THE UNIVERSITY
of EDINBURGH

Patient-led Functional Genomics Identifies Novel Drivers of Intrahepatic Cholangiocarcinoma

Nicholas Younger B.Sc (Hons)

Submitted for the degree:

Doctor of Philosophy

The University of Edinburgh

2019

Declaration

I, the undersigned, hereby declare that this thesis is solely of my own composition, and unless otherwise stated, constitutes my own work. This work has not been submitted for any other degree of professional qualification.

Nicholas Younger 14/05/2020

This work is dedicated in loving memory to:

Jane Welch

Karen Younger

Thomas Whitehill

&

_____

my kidney donor and their family.

Abstract

Intrahepatic cholangiocarcinoma (iCCA) is a rare and universally lethal malignancy arising in the liver. Incidence has steadily been increasing globally yet effective therapeutics are lacking. Currently, gold standard treatment is complete resection for which a minority of patients are suitable and recurrence rates post-resection are high. For the majority of patients unsuitable for surgery, standard chemotherapeutic options extend life by only 3-6 months on average. As such there is a pressing need to elucidate the genetic and molecular features of this malignancy so as to inform the development of more effective therapies. To date, a small but robust set of driver genes have been discovered and for some targeted therapies in development, with mixed results in clinical trials.

This work aimed to extend the set of known driver genes by exploring the long tail of infrequently mutated genes in patient sequencing data. Through computational prediction and *in vivo* screening, a set of driver genes were uncovered which cooperate with the more commonly encountered oncogenic $RAS^{G\ 2}$ isoforms of KRAS and NRAS. Validation studies *in vivo* confirmed that the membrane-cytoskeleton adaptor protein Merlin, encoded by *NF2*, and the canonical transmembrane semaphorin receptor PLXNB2 encoded by *PLXNB2* are tumour suppressors occurring at low frequency in the patient population. Mutation of these genes significantly accelerated disease progression and propagated aggressive, invasive tumour phenotypes.

Lay Summary

Intrahepatic cholangiocarcinoma (iCCA) is a rare form of cancer which occurs in the liver. For rare cancers such as iCCA the discovery of causative genes is particularly difficult. In this research the aim was to discover new causative genes in iCCA using methods that overcome this problem. In order to do this, DNA from several hundred patient tumours was analysed using a computational approach which predicts the genes which are likely to be causative based on what they do and how often they are mutated. This list of predicted genes was then used as the starting point for experimental work which tested whether these could initiate cancer by introducing these mutations in an animal system.

A single causative gene is rarely enough to cause cancer, and this is particularly true for two previously confirmed iCCA genes called KRAS and NRAS. These 2 related genes are common across many different cancer types, but in the liver they are not enough to cause cancer on their own. It was therefore valuable to determine which predicted genes can interact with these known genes to initiate cancer, and this was also investigated here

This work used patient data combined with an innovative experimental approach to uncover a set of confirmed novel driver genes in iCCA. Specifically, this included a core subset of 28 genes which cooperates with both K/NRAS mutants to cause cancer. This opens the door to the development of therapeutic approaches capable of targeting K/NRAS-mutant iCCA by targeting its co-operating driver mutations.

Acknowledgements

There are many people who have offered invaluable support throughout the last 4 years, many more than can be mentioned here. I am particularly thankful to my supervisor Dr Luke Boulter, who has provided endless guidance and advice throughout this work and beyond. This work also wouldn't have been possible without the help of the lab, particularly Scott, Mollie, and Ed, who have kept me sane and on track and often provided much needed distraction.

Thanks also need to be extended to my friends who over the last 4 years have tolerated much grumpiness and have frequently been ditched in favour of a weekend/evening experiment that 'definitely can't wait'. Sean, Stephen, Emilio, Gary, Simon, Jamie, James, Andrea, Issy, Dan, and Fiona you have all supported me through so much more than just my PhD, and my life is infinitely better for having you in it.

Lastly, and most importantly, my whole family, in particular my parents and brother, have supported me immeasurably throughout my PhD, and in particular throughout this last very challenging year.

Contents

Chapter 1 Introduction

Intrahepatic cholangiocarcinoma (iCCA) presents myriad clinical challenges and is associated with significant mortality. Tumours are asymptomatic and late-presenting and this is further complicated by marked chemotherapeutic resistance rendering most interventions ineffective. As such, the only chance for cure is complete resection or liver transplantation. This is not an option for the vast majority of patients who are unsuitable due to advanced disease featuring local invasion[2]. Of those patients who do undergo curative resection, recurrence rate is high due to primary disease commonly involving the vasculature[3] and/or lymph nodes[4 6]; indeed, even in transplant patients recurrence rates as high as 51% within the first 2 years have been reported[7]. As such there is a pressing need for novel non-surgical interventions and there is a focus in the field on molecularly targeted therapies. This has been motivated by the recent identification of recurrent targetable genomic alterations and exciting progress is continually being made in this direction[8,9]. However, the realisation of effective targeted therapies is complicated by the significant molecular and genetic heterogeneity which characterises this malignancy [0,] . In this study I aim to explore the long tail of infrequently mutated genes in iCCA in an attempt to reveal novel drivers and their modulation of common drivers. The goal of this is to expand our basic understanding of this malignancy to aid in the discovery of therapeutic targets.

1.1 Classification and Incidence of Biliary Malignancies

Cholangiocarcinomas (CCA) represent a genetically and molecularly diverse set of bile duct malignancies [2], which after hepatocellular carcinoma (HCC), constitute the second most frequent class of primary liver cancer [3]. CCAs typically arise from the ductular epithelium and can occur at any point along the length of the biliary tree. CCA is most commonly classified anatomically based on its position along the biliary tree (fig 1.1.1). As such, intrahepatic cholangiocarcinoma (iCCA) arises in the second-order bile ducts within the liver parenchyma, perihilar CCAs (pCCA) sometimes called Klatskin tumours are found around the confluence of the left and right hepatic ducts and within the common hepatic duct, and distal extrahepatic CCA (dCCA) is found within the common bile duct below the junction of the cystic and common hepatic ducts and above the Ampulla of Vater [4] (Fig. 1.1.1).

The focus of this investigation is iCCA as this is the anatomical subtype with the highest mortality despite accounting for as few as 5-10% of all CCA cases [5]. This is evidenced by the observation that in the UK iCCA has overtaken HCC is the primary cause of death from liver cancer [6]. ICCA is classified according to gross morphology as Mass-Forming (MF), Periductal Infiltrating (PI), and Intraductal Growing (IG) subtypes. Histologically there have been a number of proposed classification systems [7], but broadly speaking iCCAs are grouped into two histological subtypes which are associated with their point of origin along the biliary tree: iCCA arising

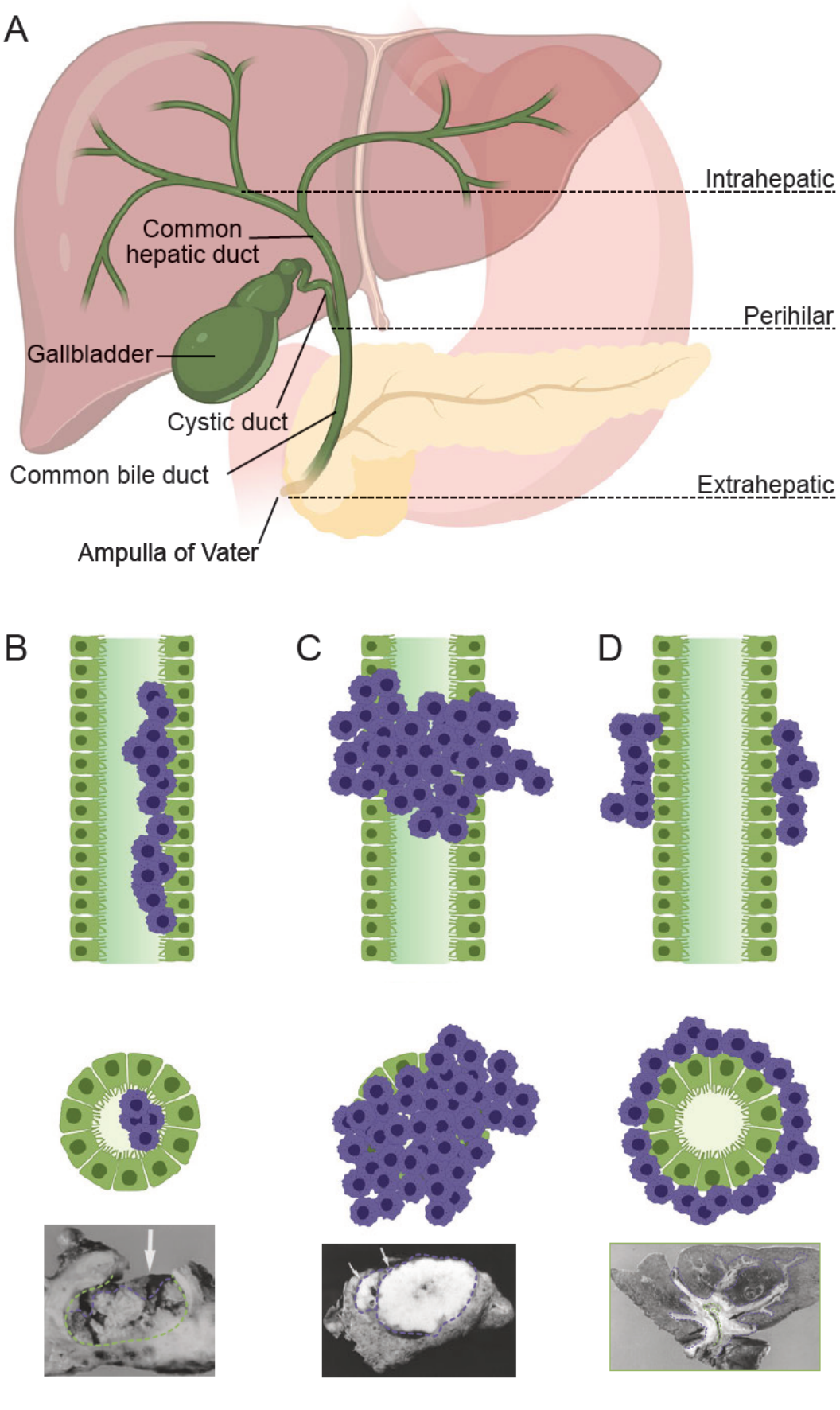

A
Intrahepatic
Common
hepatic duct
Perihilar
Gallbladder
Cystic duct
Common bile duct
Extrahepatic
Ampulla of Vater
B
C
D

Figure 1.1.1 Anatomical classification of cholangiocarcinoma: A, Anatomical classification of cholangiocarcinoma subtypes. Intrahepatic cholangiocarcinoma arises above the second-order bile ducts, i.e. above the second bifurcation of the biliary tree. Perihilar arises within the common hepatic duct up to the second bifurcation, and extrahepatic occurs in the common duct, below where the common hepatic and cystic ducts meet, and above the ampulla of Vater. B-D, Cartoon depictions of iCCA growth types with photographs showing gross morphology. B, the intraductal growing (IG) subtype describes iCCA contained within the bile ducts and spreading along the epithelial surface. As can be seen in the photograph, the bile duct wall is increased in thickness, but the tumour is contained within. C, mass-forming (MF) describes large solid masses which occlude the bile duct and spread into the parenchyma. D, periductal infiltrating (PI) subtype tumours grow around the outside of the bile ducts and track along them with moderate infiltration into the parenchyma, but without intraductal growth. Arrows indicate tumour delineated by purple dashed lines, with bile duct delineated by green dashed lines. Photographs adapted from Jae Hoon Lim, *AJR* (2003) [[8]]. Cartoon figures created with BioRender.com.

in the small peripheral ducts falls into the ‘bile ductular/mixed’ subtype, whilst iCCA arising in the large ducts are typically of the ‘bile duct/mucinous’ type [9 2] .

ICCA arises in 7 [h] decade of life in the general population, although this is modified by risk factors as discussed below[22]. The age-standardised incidence rates (ASIR) for iCCA differ from country to country but show an overall trend towards an increasing incidence globally. The rise in incidence has been deemed genuine after correcting for confounding factors such as erroneous inflation by classification artefacts. Such artefacts are attributed to differences in coding between, and time of adoption of, different editions of the International Classification of Diseases for Oncology (ICD-O). The ICD-O variably switches from classifying perihilar tumours as intrahepatic or extrahepatic between editions [23,24]. Additionally, it is not uncommon for metastasis from cancers of unknown primary (CUP) to be wrongfully diagnosed as iCCA which may further exaggerate what is, nonetheless, a genuine increase in incidence.

1.2 Epidemiology and Risk Factors

CCA has an intriguing epidemiology. Incidence shows strong geographical variation due to a combination of incompletely understood genetic risk modifiers and variable exposure to risk factors often driven by local environment[25].

ASIR per 100,000 for CCA is low in Europe (UK, 2.2; Italy, 3,4) and the American Continent (US, 1.6; Canada, 0.4), with relatively higher rates in Southeast Asia (Shanghai, China, 7.6; Busan, South Korea, 8.8), but dramatically peaks at 85.0 in North East Thailand[26]. This is worth exploring in more detail. On the surface this peak is attributable to widespread exposure to the carcinogenic liver fluke *Opisthorchis viverrini*. Liver fluke infection represents an established risk factor for iCCA with food-borne trematodes estimated to affect 40-50 million people worldwide, disproportionately focused in Southeast Asia[27]. In China and Korea the predominant carcinogenic fluke is *Clonorchis sinensis*, and *O felineus* is prevalent in Eastern Europe [28]. Returning to North East Thailand, iCCA due to *O viverrini* is exacerbated by the interplay of cultural practices and socioeconomic factors. Primary infection is often through ingestion of a traditional raw fish dish and can be treated easily with praziquantel. However, CCA incidence is more prevalent in rural regions particularly around the Mekong River Basin in the North East where the fish vector of the fluke is found [29,30]. In rural regions chronic or repeat cycles of infection are more common, contingent on factors such as lack of access to medical care and lower levels of education [3] . Work in a hamster model of *O viverrini* infestation suggests that fluke infection alone is not sufficient to induce overt iCCA

but can do so in conjunction with other tumour-promoting exposures[32,33]. How well this experimental observation translates to humans is difficult to resolve. It has been suggested that in rural Thailand increased globalisation and a shift from agriculture to industry has exacerbated economic disparities leading to greater levels of poverty[34 36]. Some effects of poverty which increase cancer risk include higher likelihood of alcohol abuse, higher incidence of smoking, and chronic stress and these are all thought to contribute alongside fluke exposure to the increased incidence of iCCA seen in these poorer regions [37]. Subsequently CCA generally and iCCA specifically are sometimes considered diseases of poverty in Thailand[38].

In Europe, North America, and Canada the aetiology is less clear. In many cases iCCA is thought to arise sporadically with no known underlying risk factor, indeed 50% of iCCA cases are of unknown aetiology. The common pathology between known (non-genetic) risk factors is a degree of chronic hepatic and/or biliary inflammation. Inflammation subsequent to choledochal cysts, primary sclerosing cholangitis, Caroli's disease, cholelithiasis, and cirrhosis are all strongly or very strongly associated with iCCA. Hepatitis B and C viral infections are moderately or strongly associated, and more classical risk factors for cancer such as alcohol consumption, obesity, and cigarette smoking are weakly or moderately associated. This points to a prominent role of chronic inflammation as an iCCA risk modifier. In addition to inflammation, bile stasis and locally high bile acid concentrations subsequent to ductal dilation seen in cysts, Caroli's disease, and lithiasis results may have direct or indirect transformative effects on the biliary epithelium[39 43]. Whilst the above cystic diseases have a strong association with iCCA, they have

higher odds ratios for extrahepatic biliary malignancy. The clearest risk factor for iCCA is primary sclerosing cholangitis (PSC), an autoimmune disease manifesting in the biliary tree: individuals with PSC have a 400-fold higher risk of iCCA than the general population.

In the West, the exposures linked to iCCA development are still to be fully elucidated and causal mechanisms remain incompletely understood. Whether or not there is a link between heterogeneity in mutational spectra and risk exposure remains an unanswered question. It is conceivable that where a very strong risk modifier such as PSC or *O viverrini* underpins iCCA development the mutational spectra may be more uniform, whereas in 'spontaneous' disease where malignancy may arise on the background of multiple exposure to low-risk sources such as cigarette smoking, alcohol consumption, transient hepatic inflammation, etc. this heterogeneity in risk exposure may be reflected in heterogeneity of the mutational spectrum.

Germline genetic risk factors are less understood due to a paucity of genome wide association studies (GWAS) in CCA. To date only one such study is underway and it is still in its early stages. Most GWAS focus on PSC and PSC-derived CCA for which some risk alleles are known but these do not translate to non-PSC CCA. This is exemplified by the finding that two single nucleotide variants in the natural killer cell receptor G2D (NKG2D) gene are associated with an increased risk of CCA development in patients with PSC [44], but this did not translate on independent validation to sporadic CCA [45].

Overall, risk factors for iCCA are many and varied and this is reflected in a geographically heterogeneous epidemiology and a genetically heterogeneous mutational landscape, as will be discussed.

1.3 Cellular, Molecular, and Genetic Pathogenesis of iCCA

1.3.1 Cell of Origin in iCCA

Primary liver malignancies are increasingly thought of as a continuum where the tumour phenotype is a product of the cell of origin, the mutational landscape, the microenvironment, and exogenous exposures[46]. This is supported by shared molecular features between HCC and iCCA[47 49], the existence of rare mixed HCC-iCCA tumours, and the observation that iCCA can arise from mature hepatocytes, mature cholangiocytes[50] and bi-potent progenitor cells. The possibility of multiple cell-types of origin for iCCA can be better understood by first considering liver cell plasticity more generally. The liver has a remarkable regenerative capacity facilitated primarily by the plasticity of both the hepatocyte and cholangiocyte lineages which can trans-differentiate into each other, as well as facultative bi-potent progenitor cells the oval cells. Liver regeneration partly involves the redeployment of developmental programs resulting in the transient acquisition of stem-like phenotypes, de/trans-differentiation, and increased proliferation. However, unlike in organogenesis, this occurs in the context of a pro-inflammatory milieu and involves cells which may have many decades of accumulated mutations. Repeated or chronic activation of regenerative programs increases the risk of primary liver cancers, which may be more likely to arise from different cells of origin dependant on the type and duration of injury. At early stages, damage of the parenchyma results in regeneration from hepatocytes in early disease, but from mature cholangiocytes or progenitor cells in late-stage cirrhotic disease[5] . Generally speaking, injury that targets the parenchyma preferentially induces

hyperplasia/regeneration of mature hepatocyte origin. Only when damage is extensive, and hepatocytes enter senescence or otherwise exhaust their replicative/regenerative potential is hepatocyte replacement from progenitor cells or mature cholangiocytes possible. Chronic bile duct injury through fluke infestation, cholestasis, or primary sclerosing cholangitis on the other hand exposes the mature cholangiocyte to prolonged inflammation and carcinogenic agents (e.g. reactive oxygen species, high concentrations of bile acids). Biliary damage also induces periportal proliferation of the ducts termed a ductular reaction which has also been shown to derive from bi-potent stem cells, mature cholangiocytes and mature hepatocytes. A plethora of observations support the plasticity of the main cell types of the liver which in turn lends credence to multiple cell-types of origin in primary liver cancers. ICCA as a product of the aberrant activation of developmental programs in the context of tissue injury is further supported by the molecular mechanism driving iCCA tumourigenesis. Indeed, many of the pathways involved in development, regeneration, disease, and malignancy are shared amongst these processes.

Whilst many molecular pathways are implicated in iCCA pathogenesis, here I focus on a limited set: Notch, Hippo, Wnt, p53, and the RAS/RAF/MEK/ERK cascade. Whilst this skips over other iCCA-relevant pathways (TGFβ[52 55], PI3K/Akt/mTOR[56 59], Hedgehog signalling[60 62], Jak/STAT[63 66]), this representative subset serves to illustrate pathway crosstalk and pleiotropy in the liver during development, regeneration, and malignancy. As such they provide a general molecular framework within which the results of the present study may be contextualised.

1.3.2 Hippo Signalling

The Hippo pathway is best characterised as a regulator of organ size during organogenesis where it balances cell proliferation with apoptosis through the integration of myriad developmental queues. As such, its core components are highly conserved across metazoan lineages[67], with upstream regulators showing species-specific differences[68]. At its core, the Hippo pathway is a kinase cascade involving the activation of firstly Ste20-like serine/threonine kinases MST1/2[69,70], which in turn activate the large tumour suppressors LATS1/2 through phosphorylation of 2 highly conserved regulatory motifs in their c-terminal catalytic domain[7] . Activated LATS1/2 serves to phosphorylate the transcriptional co-activators yes-associated protein (YAP) and transcriptional co-activator with PDZ binding motif (TAZ) which prevents their translocation to the nucleus and ultimately mediates their destruction by the proteasome[72]. In the absence of an upstream activator, the kinase pathway is inactive and YAP/TAZ translocates to the nucleus. There it binds to one of several partners[73] (RUNX1/2, TBX5, SMAD, p73, ERBB-4), but archetypically one of the TEA DNA-binding proteins (TEAD1-4) to induce pro-proliferative and anti-apoptotic transcriptional programmes[74,75]. Upstream activators of the pathway include G-protein coupled receptors[76], TAO kinases[77], FAT cadherins[78], MAP4K[79], and FERM-domain containing proteins[80].

Notable in the liver is the FERM-domain containing protein Merlin[8 ,82], encoded by neurofibromin 2 (NF2), a tumour suppressor linked with the familial cancer syndrome Neurofibromatosis Type 2[83]. During liver development and regeneration Hippo signalling is required for control of organ size[84 86], with ectopic activation of

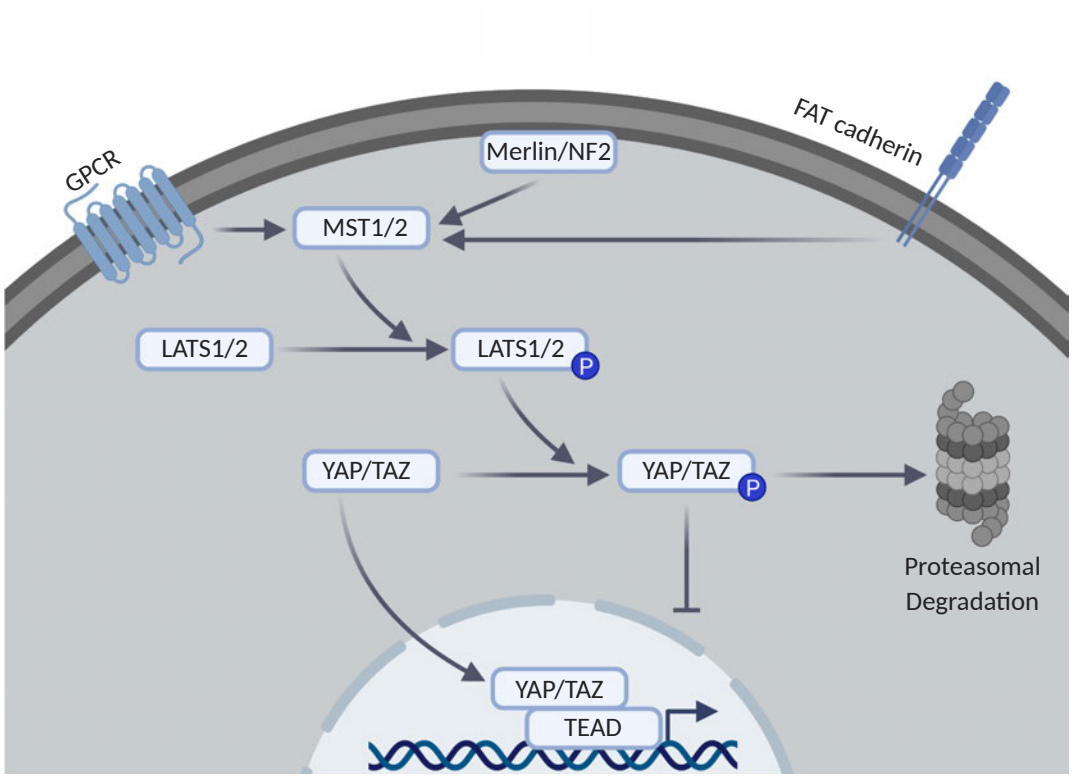

FAT cadherin
GPCR
Merlin/NF2
MST1/2
LATS1/2
LATS1/2
P
YAP/TAZ
YAP/TAZ
P
Proteasomal
Degradation
YAP/TAZ
TEAD

Figure 1.3.2.1 Hippo Pathway. The core Hippo pathway consists of a kinase cascade whereby, upon activation, MST1/2 and LATS1/2 are phosphorylated. This results in phosphorylation of YAP/TAZ which leads to its destruction by the proteasome. In the absence of activation YAP/TAZ remains unphosphorylated and is able to translocate into the nucleus where it induces transcription of target genes. Figure created with BioRender.com.

YAP or loss of NF2 causing extrahepatic choledochal cysts[87], hepatomegaly and tumour formation[88]. In the adult liver, nuclear YAP is restricted to the biliary lineage and Hippo pathway activation maintains mature hepatocyte identity by excluding YAP from the nucleus. Subsequently, nuclear translocation of YAP in hepatocytes through loss of NF2 causes dedifferentiation of hepatocytes into a progenitor state and subsequent reprogramming into a ductal-like state mediated by TEAD4 induced transcription of the cholangiocyte lineage genes NOTCH2 and SOX9[89]. Although experimental evidence exists for a role of Nf2 loss in human iCCA, the literature is sparse. In one study of human iCCA tissue sections, low expression of Merlin was correlated with poorer differentiation[90]. Loss of NF2 in mature cholangiocytes results in moderate ductal hyperplasia, but not acquisition of a progenitor cell fate or a ductular reaction.

1.3.3 Notch Signalling

The Notch pathway functions through cell-cell contact between ligand and receptor-expressing cells to elicit highly context-dependent responses. Despite the pleiotropic nature of Notch signalling, its core components are comparatively simple[9] . The core signalling module begins when a member of the single-pass transmembrane Notch receptor family (NOTCH1-4) expressed on the surface of one cell is activated through binding of any of its five Delta-Serrate-Lag (DSL) family ligands (JAG1, JAG2, DLL1, DLL3, DLL4), expressed on the surface of an adjacent cell. Binding induces a conformational change in the Notch receptor that exposes an extracellular cleavage site which is cleaved by ADAM10/17 metalloproteinases and disintegrin. A subsequent proteolytic cleavage of the

intramembrane domain mediated by γ-secretase releases the intracellular component of the receptor the Notch intracellular domain (NICD) that translocates to the nucleus. In the nucleus NICD binds the RBPJ transcription factor through displacement of co-repressors; successive binding of co-activators results in formation of the RBPJ transcriptional complex that activates Notch target genes. The most prominent Notch target genes are the transcriptional repressors hairy/enhancer of split (HES) and HES-related with YRPW motif (HEY). HES and HEY predominantly inhibit transcription of numerous target genes.

Notch signalling is pervasive in development with roles in effectively all major organ systems. One prominent role of Notch with relevance here is in the control of cell fate decisions and the establishment of tissue boundaries through lateral inhibition or induction. During liver organogenesis, Notch signalling first specifies the ductal plate the layer of cells which goes on to form the biliary tree - through interaction of Notch-expressing hepatoblasts and adjacent Jag-1-expressing portal mesenchyme; Jag-1 expression subsequently shifts to the parenchymal side and induces ductular morphogenesis in concert with other signalling pathways[92] notably TGBβ[93], which itself has been shown capable of mediating complete reconstitution of a biliary tree in the mature mouse liver[94]. Attesting to the fundamental importance of Notch in biliary cell fate decisions, germline loss of NOTCH2 or JAG1 genes results in Alagille syndrome which is characterised by the complete absence or a paucity of the biliary tree[95]. Out-with development, Notch1 and Jag1 are both expressed during liver regeneration following partial hepatectomy in rats[96]. Although this does not seem to translate directly to mice,

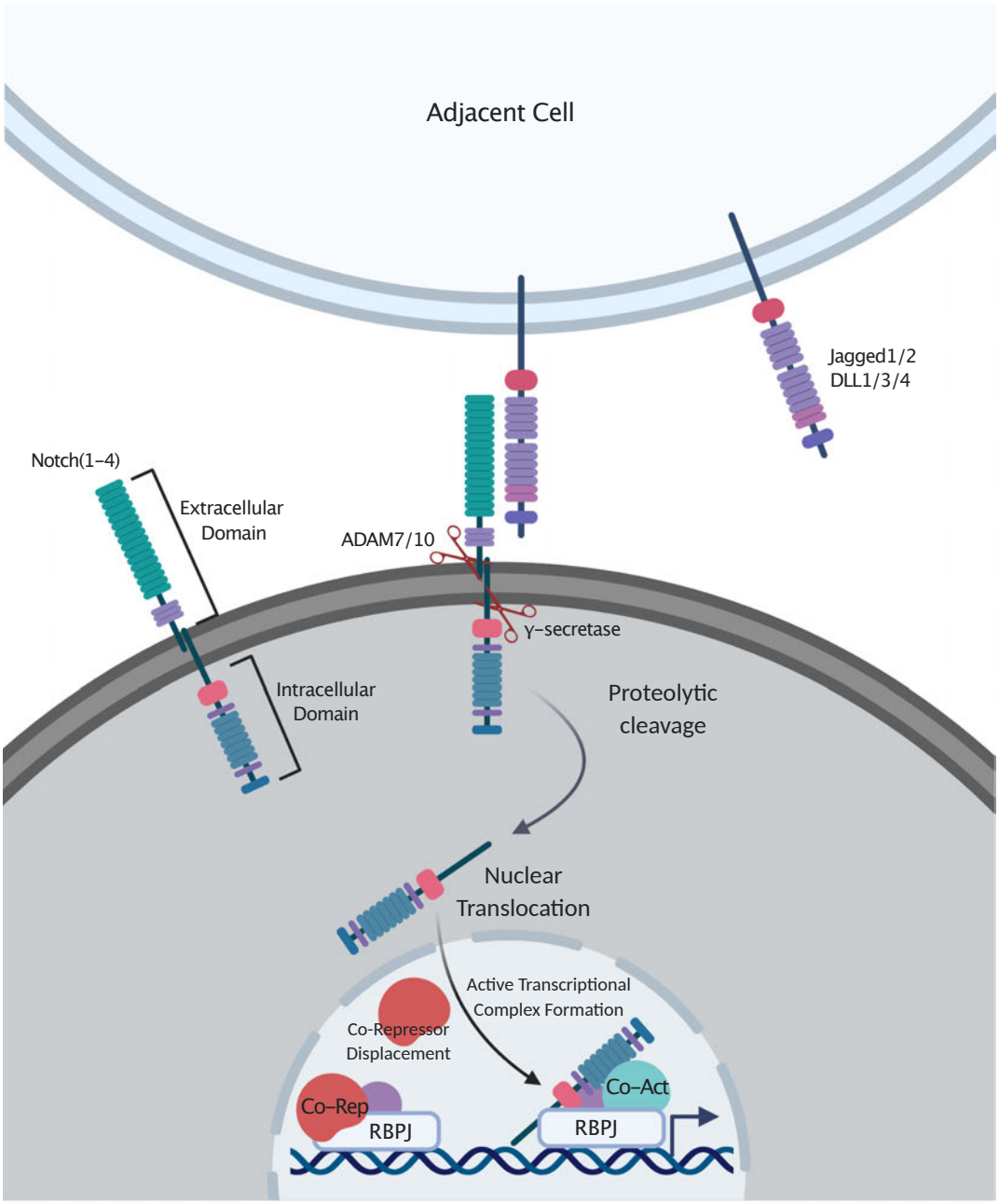

Adjacent Cell
Jagged1/2
DLL1/3/4
Notch(1–4)
Extracellular
Domain
ADAM7/10
γ–secretase
Intracellular
Domain
Proteolytic
cleavage
Nuclear
Translocation
Active Transcriptional
Complex Formation
Co-Repressor
Displacement
Co–Rep
RBPJ
Co–Act
RBPJ

Figure 1.3.3.1 Notch Pathway. The Notch pathway is a contact-dependent pathway whereby a ligand-receptor pair is formed between 2 adjacent cells. In the receiving cell this leads to successive proteolytic cleavage of Notch receptor peri-membrane domains liberating the intracellular domain, NICD. This translocates to the nucleus where it displaces TEAD co-repressors and initiates formation of a co-activation complex to drive transcription of target genes. Figure created with BioRender.com.

conditional loss of NOTCH1 induces nodular regenerative hyperplasia through the continuous proliferation of hepatocytes[97] with concomitant vascular remodelling and without fibrosis[98].

In line with its role in specifying cell fate in the liver, Notch-mediated transdifferentiation of mature hepatocytes into iCCA has been shown in two complementary mouse models. Forced activation of downstream effectors through expression of NICD in mature hepatocytes causes the formation of cystic cholagiocellular lesions[99]. When this is combined with overexpression of Akt the result is differentiated, ductular lesions with numerous characteristics of human iCCA[99]. This was also observed in a second model which used the TAA liver damage diet as a pro-oncogenic stimulus in place of Akt, and further showed that loss of the Notch effector Hes1 abrogated this phenotype [00]. In early HCC the transcription factor RUNX3 acts to limit HCC progression by directly interacting with NICD to suppress Notch signalling [0] . When not suppressed in HCC, Notch2 signalling produces poorly differentiated tumours and correlates with immature morphology, aggressiveness and advanced clinical stage in human samples [02]. This observation further supports the role of aberrant Notch signalling in hepatocyte lineage plasticity. Finally, O'Rourke *et al.* [03] comprehensively profiled the role of NOTCH signalling in human iCCA, and delineated a transcriptomic signature for prospective stratification of iCCA patients into groups likely to respond favourably to pan-γ-secretase inhibitors. Importantly, they demonstrated a pivotal role for stromal-epithelial NOTCH signalling as an important modality in human iCCA. They further showed the simultaneous upregulation of multiple Notch receptors and

ligands thus necessitating the requirement of an inhibitor targeting universal components of the pathway (i.e. γ-secretase) as opposed to those targeting individual ligands/receptors. This study supports and elaborates on a previous study by Guest *et al.* [04] demonstrating differential overexpression of NOTCH3 in human sections, as well as in transgenic and chemical-induced models of iCCA in mice and rats, respectively. In this study the authors showed that tumour survival was mediated by activation of the PI3k/Akt pathway, which represents a druggable target in iCCA[56, 05, 06].

1.3.4 The Wnt pathway

The canonical and non-canonical Wnt signalling pathways constitute a complex network with numerous ligands and extensive crosstalk with each other [07 09] and many other pathways [0,].

Canonical Wnt signalling is fundamental as a regulator of proliferation and stemness in both the adult and the embryo. The core components of the canonical pathway revolve around the multi-function protein β-catenin [2]. In one capacity it acts as a plasma membrane-associated linker, bridging the interaction of transmembrane E-cadherin to cytoskeletal actin filaments to mediate cell-adhesion and mechanotransduction [3]. This function is critical in gastrulation and is conserved in multicellular organisms, preceding the divergence of Cnidaria and Bilateria [4, 5]. In another capacity, β-catenin acts as a transcriptional regulator through regulated [6] interaction with the T-cell factor (TCF)/lymphoid enhancer binding factor (LEF) family of transcription factors [7, 8]. In the absence of pathway

activation β-catenin not involved in cell-adhesion is turned over by the destruction complex a dynamic assemblage with core components consisting of the adenomatous polyposis coli (APC) protein, casein kinase-1α (CK1α), glycogen synthase kinase3β (GSK3β), and the scaffold protein Axin [9]. Additionally, protein phosphatase A2 (PPA2) is also associated with the complex although its function is not definitively resolved [20]. Phosphorylation of β-catenin allows its recognition and binding by the E3-ubiquitin ligase β-TRCP-SCF complex [8, 9, 2] which ultimately results in its destruction by the proteasome. Pathway activation is initiated when any of the family of canonical ligands secreted, lipid-modified, glycoproteins [22] binds to the seven-pass transmembrane co-receptor Fzd. This induces heterodimerization of activated Fzd with either of its binding partners, the LRP5/6 co-receptors, and produces the active Fzd-LRP5/6 receptor complex. GSK3β-mediated phosphorylation of the active LRP5/6 results in sequestration of Axin to the membrane through interaction with the scaffold protein, Dishevelled. This ultimately sequesters the destruction complex to the plasma membrane, allowing β-catenin to accumulate in the cytoplasm and translocate to the nucleus where it activates target genes through interactions with its binding partners primarily through a conserved armadillo repeat domain, crucial for many of β-catenin's protein interactions, and a therapeutic target [8].

Non-canonical forms of Wnt signalling are β-catenin-TCF/LEF-independent but are still mediated by the cognate Fzd receptor in complex with numerous co-receptors. The best characterised of these alternative pathways are the Wnt/planar-cell polarity (PCP) and Wnt-Ca^{2+} pathways. In its simplest incarnation, the former acts

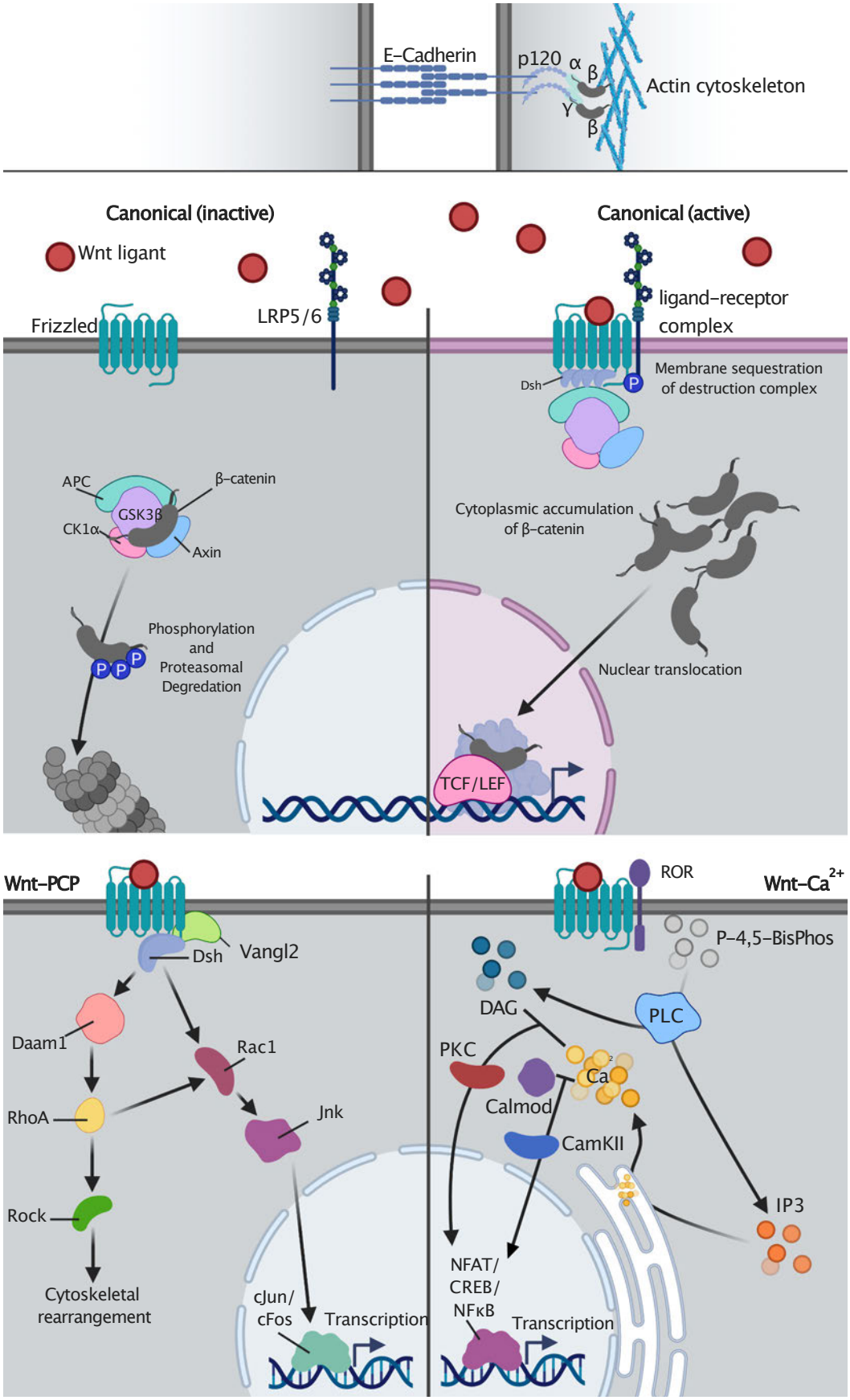

E-Cadherin
p120
α
β
γ
β
Actin cytoskeleton
Canonical (inactive)
Canonical (active)
Wnt ligant
Frizzled
LRP5/6
ligand–receptor complex
Dsh
P
Membrane sequestration of destruction complex
APC
β–catenin
GSK3β
CK1α
Axin
Cytoplasmic accumulation of β–catenin
Phosphorylation and Proteasomal Degredation
Nuclear translocation
TCF/LEF
Wnt–PCP
Wnt–Ca2+
ROR
Dsh
Vangl2
P–4,5–BisPhos
DAG
PLC
Daam1
Rac1
PKC
Ca2
Calmod
RhoA
Jnk
CamKII
IP3
Rock
Cytoskeletal rearrangement
cJun/ cFos
Transcription
NFAT/ CREB/ NFκB
Transcription

Figure 1.3.4.1 Wnt Pathways. A, β-catenin associates with alpha and gamma catenins to bridge the interaction of E-cadherin and the actin cytoskeleton at cell-cell junctions. B, in the absence of ligand, inactive canonical Wnt signalling involves the constant turnover of β-catenin by the destruction complex. Upon ligand binding and heterotypic receptor dimerisation the destruction complex is sequestered to the membrane. This inhibits the destruction of β-catenin, allowing it to accumulate and translocate into the nucleus where it induces transcription of target genes. C, The Wnt-PCP pathway is β-catenin-independent and is activated when Frizzled and Vangl2 form a receptor dimer upon binding of a non-canonical Wnt ligand. This can result in both transcription of targets via Jnk/Jun/Fos cascade or through cell polarisation through rearrangement of the cytoskeleton mediated by RhoA/ROCK. D, The Wnt-Ca^{2+} pathway involves the production of second messengers (DAG, IP3) by phospholipase C upon activation of the ROR receptor. This induces intracellular Ca^{2+} release from the endoplasmic reticulum and subsequent calcium-dependent activation of CamKII and PKC. This leads to activation of numerous transcription factors. Figure created with BioRender.com.

via the Fzd/Vangl2 receptor complex resulting in recruitment and activation of Dishevelled which subsequently releases inhibition of DAAM1. This initiates a RAC1/RHOA/ROCK/JNK cascade that results in actin cytoskeletal rearrangements and c-Jun/c-Fos transcription factor activation in the nucleus. This core pathway establishes cell polarity in mesenchyme and epithelial sheets, and is therefore indispensable in embryogenesis and organogenesis where it mediates convergent extension during gastrulation and neurulation [23, 24], and establishes tissue polarity [25] and left-right asymmetry [26]. In Wnt-Ca^{2+} signalling, Fzd/Ror1/2 binding initiates phospholipase C (PLC) activation which catalyses the production of inositol-1,4,5-triphosphate (IP3) and 1,2-diacylglycerol (DAG) from membrane-bound phosphatidylinositol-4,5-bisphosphate. IP3 induces Ca^{2+} release from the endoplasmic reticulum which, in combination with calmodulin, activates calcium-calmodulin-dependent protein kinase II (CaMKII). Meanwhile, DAG, in concert with the efflux of Ca^{2+} from the endoplasmic reticulum, activates protein kinase C (PKC). The resultant activation of PKC and CaMKII signals to activate transcription factors such as NFκB, CREB, and NFAT [27].

Wnt signalling in virtually all of its incarnations plays a prominent role in development, and this is no less true in liver organogenesis [28]. Following gastrulation, Wnt signalling induces the foregut [29, 30] and subsequently induces expansion of bi-potent hepatoblasts in the liver bud [30]. Importantly, Wnt acts as a key factor in the specification of biliary fate in coordination with Notch. Decaens *et al.* [3] demonstrated that forced activation of β-catenin through APC loss in the mouse embryo resulted in failure of hepatocyte specification from hepatoblasts

and the concomitant differentiation of ductular structures of biliary identity. Lineage tracing further showed that APC-null hepatoblasts formed mature bile ducts when transplanted into the adult liver but did not contribute hepatocytes. Beside from a clear role of canonical Wnt signalling in biliary specification, Cui *et al* [32]. showed the importance of Wnt-PCP signalling in biliary fate specification and ductular morphogenesis in the zebrafish, with its loss resulting in reduced development/branching of the biliary tree and absence or incorrect sidedness of the gall bladder.

Mutations in the regulatory domain of β-catenin, required for its phosphorylation by GSK3β, are found in >50% of paediatric hepatoblastoma, as are loss-of-function mutations in APC and AXIN. To a lesser extent, mutations in these and others in the Wnt pathway are also found in HCC where they promote cell proliferation, metastasis, and aggressiveness [33]. As is the case in hepatoblasts and earlier progenitors, Wnt-signalling was shown to target putative HCC stem-like cells and promote their stemness and proliferation [34]: suppression of Wnt signalling by endogenous micro-RNAs results in loss of stemness [35]. Somewhat paradoxically given its role in biliary fate specification in the embryo, in one study Wnt signalling in early HCC was correlated with better hepatocytic differentiation that was subsequently lost through a gradual increase in hedgehog signalling with concomitant Wnt loss during the temporal progression of HCC development [36].

In human iCCA, our lab [37] previously demonstrated that WNT7B and WNT10A are expressed in human samples and that Wnt signalling can drive iCCA. This is at odds with the sparsity of Wnt pathway mutations in iCCA. The authors subsequently

resolved this disparity through discovery that inflammatory ($CD68^{+}CD163^{-}$) macrophages in the tumour stroma induce Wnt expression in iCCA epithelium through secretion of WNT7B. Subsequent loss of macrophages or pharmacological inhibition of ligand secretion by Porcupine inhibition abrogated tumour growth and caused partial regression. In addition to macrophage-derived Wnt activation, iCCA may also show increased Wnt activity as the result of decreased expression of negative Wnt regulators, such as SRC-like adaptor protein (SLAP) [38]. Although the mechanism of SLAP expression loss was not explored in the corresponding study, others have shown that iCCA frequently features epigenetic promoter silencing of other regulators of Wnt [39]. Notably, Merino-Azpitare *et al.* identified SOX17 as a novel tumour suppressor in iCCA. They showed that its loss induced Wnt-dependant proliferation (amongst other changes) through the canonical pathway. They subsequently showed that in human samples the SOX17 promoter was frequently hypermethylated, which, through decreased expression, correlated with decreased overall survival following resection.

1.3.5 The p53 tumour suppressor

The tumour suppressor p53 has been dubbed the “guardian of the genome”; compromised p53 function through mutation or dysregulation of its numerous regulatory pathways is a pervasive pan-cancer hallmark. The diversity of signals p53 can receive and transduce is a product of a highly modular domain structure rich in intrinsically disordered domains [40–43] and posttranslational modification (PTM) sites [44]. The former provides flexible binding surfaces through which a plethora of proteins can interact, and the latter provides residues which can be

dynamically modified to modulate p53 function. Further diversity in function is attributable to the fact that p53 is expressed in multiple isoforms with often antagonistic functions [45], and because the active p53 protein is a homotetramer. This adds additional opportunities for fine-tuning. For example, in unstressed cells p53 monomers, dimers, and tetramers exist at low levels, and tetramerisation upon DNA damage is reliant on co-factors and PTMs which themselves are subject to numerous layers of control [46]. Consequently, p53 functions as a network hub [47], integrating myriad intracellular and extracellular signals which in turn control cell cycle and transcriptional programs. As a tumour suppressor it initiates reparative programs and induces cell-cycle arrest upon DNA damage sensing [48], transduced by ATM, amongst others. If DNA damage is irreparable, p53 initiates apoptotic programs [49]. At its very core p53 is regulated by interactions with the MDM2 and MDMX proteins; these homologues perform non-redundant roles in modulating p53 individually, and most effectively inhibit p53 when acting in concert [50, 5] . The E3 ubiquitin ligase MDM2 is itself a transcriptional target of p53 and the resulting feedback loop [52] whereby active p53 induces MDM2 expression which subsequently ubiquitinates p53 causing its nuclear export and proteasomal degradation produces clock-like pulses of p53 activity which continue until DNA damage is resolved or the cell dies [53, 54].

Whilst the role of p53 in development is comparatively understudied, there is increasing evidence to suggest it plays a role in embryonic stem cell maintenance [55] and differentiation in a highly context-dependent manor [56] [59]. Whilst a role for p53 in liver organogenesis has not yet been investigated, there is, however, an

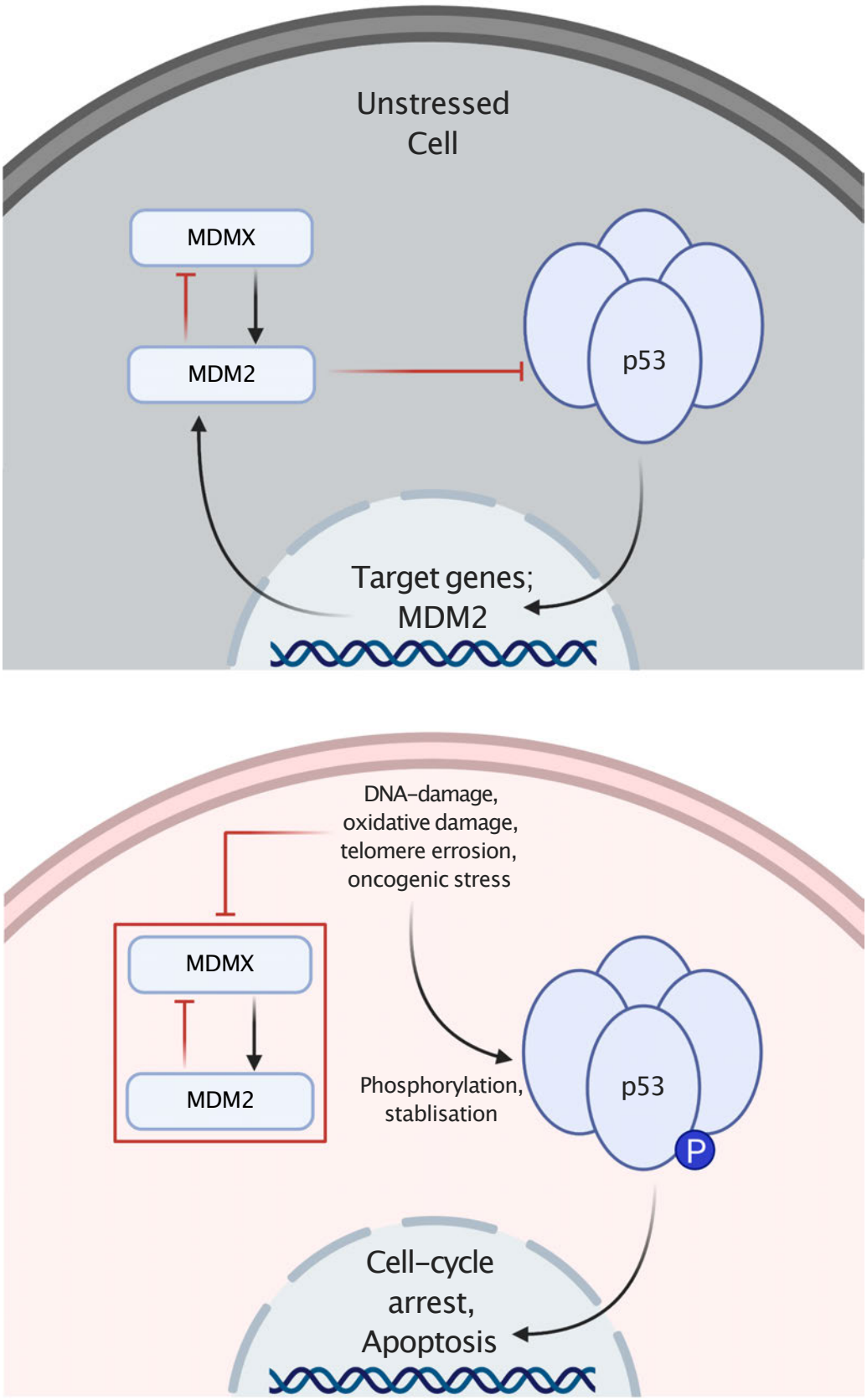

Unstressed
Cell
MDMX
MDM2
p53
Target genes;
MDM2
DNA–damage,
oxidative damage,
telomere errosion,
oncogenic stress
MDMX
MDM2
Phosphorylation,
stablisation
p53
P
Cell–cycle
arrest,
Apoptosis

Figure 1.3.5.1 p53 regulation. In unstressed cells, p53 is maintained at a low level through a feedback look involving its regulators MDM2 and MDMX. MDM2 is a transcriptional target of p53 so as p53 becomes active it upregulates MDM2 which in turn inhibits p53, maintaining its activity within a tight range. MDM2 and MDMX co-regulate each other as well as both inhibiting p53 through different mechanisms, thus leading to a fine-tuned control of resting p53 levels which can be modified through regulation of both MDM2 and MDMX. Upon cell stress the MDM2/MDMX proteins are inhibited through myriad stress-specific mechanisms and p53 is stabilised through phosphorylation. It then translocates into the nucleus to initiate cell-survival, cell-cycle arrest, or apoptotic transcriptional programmes. Figure created with BioRender.com.

emerging role for p53 in various liver pathologies and regeneration [60] [62]. For example, Borude *et al.* [63] demonstrated significantly greater levels of acetaminophen-induced damage in Trp53-null mice. Nonetheless, despite this increased damage and delayed the onset of regeneration, Trp53-null mice resolved injury in a time comparable with WT mice due to increased levels of regeneration attributable to faster cell-cycle progression.

TP53 occupies a place amongst the handful of frequently mutated genes in iCCA [64] [66]. A study by Ong *et al.* [67] and its follow-up by Chan-on *et al.* [68] indicate that TP53 mutations are enriched in fluke-related iCCA. The causal mechanisms which might favour TP53 mutation in fluke-related iCCA remains to be elucidated. One potential causal link may be the relationship between mutational processes active on tumour genomes and selection for driver mutations. Temko *et al.* [69] demonstrated that different mutations arise in driver genes as a result of mutational processes active on the genome, and that in some cases these are differentially selected. As such, the mutational processes active as a result of fluke infestation and chronic inflammatory processes warrant further investigation in iCCA; particularly in light of the observation that TP53 in the liver preferentially harboured $TP53^{R249S}$ mutations as a consequence of aflatoxin exposure [69] (accounting for 16% of p53 mutations in the cohort used in this study). Several studies also report a statistical relationship between PSC and TP53 mutations in iCCA whereby TP53 mutation is not detected in PSC samples but is present in iCCA which has developed on a background of PSC [70] [72], suggesting that TP53 loss is a precipitating event mediating the transformation of PSC into iCCA. Loss of cell identity and lineage

plasticity is proven to be one consequence of p53 loss. Experimental evidence shows that TP53 loss in combination with KRAS gain-of-function mutations [73, 74] or NICD expression [75] drives iCCA of both cholangiocyte and hepatocyte origin; the former progressing through a pre-malignant biliary intraepithelial neoplastic (BilIN) sequence prior to iCCA formation, and the latter emerging directly from hepatocytes with concomitant lineage switching to iCCA. In line with these observations, Tschaharganeh *et al.* [76] demonstrated that wild-type p53 acts as a gatekeeper of cellular plasticity in mature hepatocytes, and that its loss caused aggressive, dedifferentiated tumours, formation of which was accelerated in conjunction with Yap expression. The authors showed that the stem cell marker Nestin functionally participated in the acquisition of stemness and tumour progression, and that it is ordinarily repressed by wild-type p53 indirectly through binding with the Sp1 transcription factor. Poorly differentiated tumours which arose from p53 loss and Yap expression could subsequently be directed towards an HCC or iCCA lineage through expression of Wnt or Notch pathways, respectively.

1.3.6 The RAS/RAF/MEK/ERK kinase axis

The final discussion of pathways centres around the small GTPase superfamily members KRAS and NRAS. The small GTPases catalyse the hydrolysis of guanosine triphosphate (GTP) to guanosine diphosphate (GDP) and make up a family of more than 150 members, of which KRAS, NRAS, and HRAS are the archetypal founders [77]. The RAS proteins act downstream of receptor tyrosine kinases to transduce extracellular signals, and in this capacity they function as a binary switch. In their

GDP-bound state they are ‘off’ and do not engage their effectors, but upon binding of GTP they shift to an active ‘on’ conformation. The cycle of GTP binding, GTP hydrolysis, GDP release, and GTP re-binding occurs at too slow a pace to be physiologically relevant. Therefore, despite their intrinsic GTPase activity, GTP catalysis is aided by interaction with GTPase activating proteins (GAPs). Relatedly, the release of GDP to allow recharging is mediated by guanine nucleotide exchange factors (GEFs) [78]. Upon RTK activation (e.g. EGF binding EGFR), a receptor complex forms involving the adaptor protein growth factor receptor bound 2 (GRB2) which is constitutively bound to Son of Sevenless (SOS), a GEF [77]. SOS and RAS-GDP/GTP reciprocally regulate each other [79, 80], and SOS mediates the release of GDP allowing RAS activation through binding of GTP, mediated by GAPs such as p120-RasGAP [8] or NF1 [82].

In its active conformation, RAS initiates a kinase cascade whereby sequential kinases are activated through phosphorylation. Activated RAS phosphorylates RAF a serine/threonine kinase which results in its dimerisation and activation. Activated RAF subsequently activates MEK which activates the terminal effector, ERK. Through translocation to the nucleus ERK engages transcription factors to induce a diverse range of cell-physiological changes including proliferation, differentiation, and survival. Activated ERK can also be retained in the cytosol where it potentiates the activation of pro-apoptotic proteins [83]. Aside from the RAF/MEK/ERK axis, Active RAS can also signal through the PI3K/AKT pathway, the PLC/Ca^{2+}/PKC pathway, and others [84, 85].

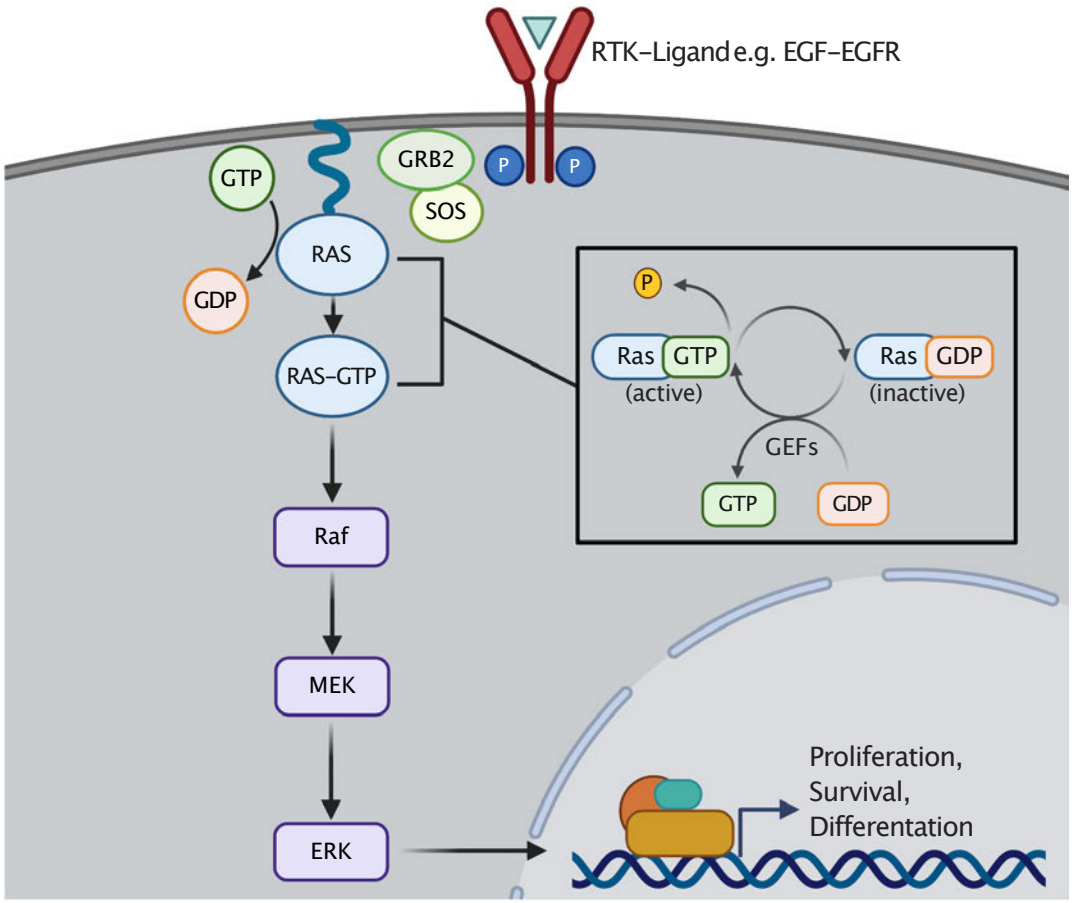

RTK–Ligand e.g. EGF–EGFR
GTP
GRB2
P
P
SOS
RAS
GDP
RAS–GTP
P
Ras GTP
(active)
Ras GDP
(inactive)
GEFs
GTP
GDP
Raf
MEK
ERK
Proliferation,
Survival,
Differentation

Figure 1.3.6.1 RAS/RAF/MEK/ERK Cascade. Upon receptor tyrosine kinase activation, RAS is switched from its inactive (GDP-bound) state and into its active (GTP-bound) state, mediated by SOS, a GEF. In the active conformation, a cascade is initiated involving successive phosphorylation of RAF, MEK, and ERK, ultimately leading to the induction of proliferative and pro-survival programmes. Figure created with BioRender.com.

Oncogenic substitutions at the G12, G13, and Q61 amino acid residues abrogate GTP->GDP catalysis by preventing *in trans* insertion of the catalytic arginine finger provided by GAPs [86]. This results in a constitutively 'on' RAS protein which perpetuates downstream kinase pathway activation leading to pro-proliferative oncogenic phenotypes.

Oncogenic K-, N- and HRAS are found in a broad spectrum of cancers, although they do show cancer specificity. For example, oncogenic KRAS occurs in almost 100% of pancreatic ductal adenocarcinoma [87, 88], whilst oncogenic HRAS is the predominant RAS oncogene in head and neck cancer [89, 90]. The reason for cancer specificity has been extensively investigated. Notable contributions include work by Haigis *et al.* [9] who showed that oncogenic KRAS and NRAS confer different phenotypes in the colon. In their study, KRAS caused expansion of stem cells, whilst NRAS did not effect growth but did confer resistance to apoptosis. This may explain preferential selection for KRAS mutants which in turn may explain the relatively increased frequency of oncogenic KRAS compared to NRAS in colorectal cancer. In an elegant study by To *et al.* [92] the authors showed that despite the predominance of oncogenic KRAS in lung cancer, oncogenic HRAS could fulfil its role when expressed from the endogenous KRAS locus but not from the endogenous HRAS locus. This suggested that in the lung at least, the gene-regulatory landscape is a prevailing factor in RAS isoform preference.

Both oncogenic KRAS and NRAS are found in iCCA but have different frequencies and different oncogenicities [93], with $KRAS^{G\ 2}$ mutations occurring more frequently and causing more aggressive phenotypes. In iCCA, codon changes targeting the G12,

G13, and Q61 residues are found at various frequencies. A study by Chung *et al.* [93] assessing the differential oncogenicity of oncogenic RAS isoforms in the liver found that the main modifier of oncogenicity was residue. $KRAS^{G12}$ expression in the liver produced tumours with much higher penetrance than $NRAS^{G12}$, but both $HRAS^{Q61}$ and $NRAS^{Q61}$ mutations were comparable with $KRAS^{G12}$. The reasons for different potencies at different oncogenic residues is incompletely understood [94, 95] but are thought to be cell-specific.

1.3.7 Molecular Pathway Crosstalk

Reconciling the vast and occasionally contradictory evidence in the literature on pathway dysregulation in iCCA to produce a unified model is a formidable challenge, and further complexity is added due to the fact that the pathways above do not exist in isolation. Pathway crosstalk is a common feature of malignancies and here select examples will be used to illustrate the importance of considering this.

Osteopontin, a multifunctional bone sialoprotein, is a critical inflammatory regulator in liver disease [96–98] and regeneration [99], with promiscuous roles in cancer through interaction with a versatile complement of receptors[200]. Recently, Zheng *et al.*[20] demonstrated that osteopontin was elevated in iCCA and correlated strongly with poor survival and recurrence due to metastasis. They went on to show that osteopontin recruits ERK2 which phosphorylates β-catenin at S657 and increases its stability, thus allowing it to accumulate in the cytoplasm and translocate to the nucleus. Here then we see the interaction of kinase signalling

through MEK-ERK with the canonical Wnt effector, resulting in its activity even in the presence of a functional destruction complex/the absence of Wnt signalling. Staying with Wnt, Wnt-Hippo crosstalk is seen vertically at multiple points along the respective pathways. At the 'top' the Hippo inducer Merlin (NF2) interacts with the canonical Wnt co-receptor LRP6 to limit its sensitivity. When Wnt signalling is below a certain threshold Merlin binds LRP6 and inhibits Wnt signalling. Frizzled binding by Wnt3a results in an increase of intracellular phosphatidylinositol-4,5-bisphosphate which indirectly induces phosphorylation of Merlin. This causes its release from LRP6, thus activating canonical Wnt signalling only above a certain threshold of Wnt ligand binding. Subsequently, loss of NF2 in cancer increases the sensitivity of Wnt signalling by lowering its threshold for activation[202]. At the 'bottom', the β-catenin/LCF4 transcriptional factor complex binds to an enhancer in the first intron of the Hippo effector YAP, driving its expression and resulting in increased downstream Hippo signalling[203]. Additionally, nuclear translocation of YAP and β-catenin are jointly inhibited by STK11 itself recently predicted as a CC driver[204] the expression of which is correlated with increased survival in gastric cancer[205].

Finally, as discussed, Notch activity causes de-differentiation of hepatocytes and is a pervasive and druggable feature in iCCA. Several different mechanisms have been described whereby activation of Notch signalling represses p53[206 208]. In one proposed mechanism Notch signalling indirectly targets the ARF-mdm2-p53 tumour surveillance axis, resulting in mdm2 activation and subsequent p53 repression. In this system inhibiting the mdm2-p53 interaction restored p53

activity and allowed for tumour regression. Considering the substantial, but incomplete penetrance of TP53 mutations in iCCA and the pervasive activation of Notch in certain patient subgroups, this may represent a synergistic therapeutic approach alongside γ-secretase inhibitors where patients harbour wild-type TP53.

1.3.8 Genetics

The mutagenic processes acting on cells over a lifetime are many and varied, and this is reflected by at least some degree of genetic heterogeneity observed across all cancer types. However, such heterogeneity is pronounced in iCCA where it is amongst its most defining features. As such, iCCA shows extensive low-penetrance drivers distributed across diverse signalling pathways[209], and complete resolution of this set remains a challenge. The root cause of this heterogeneity is incompletely resolved but is likely, at least in part, due to the varied exposure of the liver to toxic insult throughout life. There is ample opportunity for diverse mutagenic exposure due to its role in xenobiotic metabolism[20,24], as well as its targeting by oncogenic organisms (HBV/HCV[25,27], flukes[28,28,22]) and exposure to inflammation in chronic disease resulting in repeat damage/regeneration cycles.

In iCCA, exome and targeted sequencing studies [64, 67, 68,209,27,222,223] have uncovered a set of recurrently mutated genes that are consistent across cohorts: TP53, KRAS, IDH1/2, PIK3CA, and FGFR2-fusions feature most prominently. A propensity towards drivers affecting chromatin metabolism ARID1A, ARID2, BAP1, PBRM1 are also observed, and others have shown recurrent deletion or putative loss-of-

function mutations in the cell cycle regulator phosphatase and tensin homologue (PTEN) and the related cell-cycle regulators CDKN2A/B.

The largest study to date by Nakamura *et al.* [65] contributed substantially to this effort by delineating targetable alterations which segregate by anatomical location

previous studies largely grouped anatomical subtypes under the umbrella of 'cholangiocarcinoma'. This study showed that IDH1/2, BAP1, and FGFR1/2/3 mutations/fusions were preferentially found in iCCA. Amongst those alterations which are targetable IDH1/2 and FGFR amplifications/fusions are of particular note.

Fibroblast growth factor receptors (FGFRs) are a family of 5 receptor tyrosine kinases which interact with a family of 18 ligands. Activation of FGFR induces signalling primarily through the RAS/MAPK and the PI3K/Akt pathways, and in some cases activates PKC through a mechanism similar to that seen in Wnt/Ca^{2+} signalling[224]. In iCCA, amplifications of FGFR genes are observed, but more common are FGFR2 fusions. Various fusion partners have been found (*BICC1* and *AHCYL1*[225], *PPHLN1*[222], *CCDC6*[226], *KCTD1* and *TXLNA* [65]) and these universally preserve the FGFR2 kinase domain. Experimental validation in Nakamura *et al.* [65] and Arai *et al.*[225] showed ligand-independent autophosphorylation of the kinase domain activation loop, resulting in increased MAPK signalling and oncogenic phenotypes. FGFR2 fusions show mutual exclusivity with oncogenic RAS mutations demonstrating that activation of the downstream MAPK pathway is the likely target of selection and duel aberrations confer no further selective advantage. Sia *et al.* suggests that these fusion events are potentially caused by downregulation of the

DNA-damage response genes ATM and ATR which were found to be differentially downregulated in FGFR2-fusion tumours. These proteins are responsible for regulating the stability of fragile sites chromosomal regions prone to breakage during replication stress, such as that found in cancer relevant here as FGFR2 is found within such a site on chromosome 10 (FRA10F).

IDH1 and IDH2 are isocitrate dehydrogenase enzymes which function as homodimers to bi-directionally catalyse the interconversion of isocitrate and α-ketogluatarate (α-KG) during the Krebs cycle (with the forward direction strongly favoured under physiological conditions). They play central roles in several metabolic processes including modulation of redox status, glutamine metabolism and lipogenesis[227]. Point mutations affecting the R132 residue in IDH1 or the R172 and R140 residues in IDH2 map to arginine residues in the binding pocket, abrogating the ability of IDH1/2 to bind its canonical substrate. Dang *et al.*[228] showed that these mutations resulted in neomorphic activity whereby mutant IDH catalysed the conversion of α-ketoglutarate to 2-hydroxyglutarate a potent oncometabolite. This is consistent with the finding that IDH$^{R\ 23/R\ 40/R\ 72}$ mutations are monoallelic, due to the requirement of the WT product α-ketoglutarate as the substrate for the mutant IDH. 2-HG has been shown to competitively inhibit enzymes that use α-KG as co-factors. Such enzymes include epigenetic modulators responsible for DNA and histone demethylation, and in glioma this mutation alone is sufficient to extensively remodel the methylome (and by extension, the transcriptome)[229]. A critical consequence of this was reported first by Lu *et al.*[230] who showed that 2-HG produced by mutant IDH prevents histone demethylation

required for lineage specification, thus locking cells in a progenitor state. Saha *et al.*[23] extended this observation to iCCA where they showed that neomorphic IDH blocks hepatocyte differentiation and leads to an expanded pool of bi-potent progenitors primed for oncogenic transformation. This was reinforced by the identification of a hepatic progenitor transcriptional signature in IDH mutant iCCA. Analysis performed on TCGA RNA-seq data by Farshidfar and colleagues[209] identified a subset of transcriptionally segregating tumours all of which carried IDH1/2 hotspot mutations, and this was validated in an independent cohort. This group was enriched for expression of genes active in mitochondrial biosynthesis, oxidative phosphorylation, Krebs cycle, and chromatin modification.

Despite substantial progress in delineating the (targetable) driver landscape in iCCA, several key hurdles remain to be overcome. clinical trials of targeted therapies are difficult to implement due to rarity of disease and general molecular and genetic heterogeneity which hinders precise stratification. Increasingly sophisticated patient stratification frameworks [84, 85] and efforts to group patients based on key classifier mutations [66] are undoubtedly at the forefront of advancing strategies for patient-tailored therapy. With this in mind, it is frustrating to observe how often targeted therapies with tolerable toxicity show inconsistency in patient response. For example, 2 trials targeting the frequently dysregulated MAPK signalling pathway with the MEK inhibitors binimetinib[234] or selumetinib[235] showed mixed response across patients. Patients harbouring KRAS mutations that would be expected to benefit from such targeted therapy experiencing no or unsustained response, whilst some patients with unknown alterations showed

benefit. This is at odds with efficacious preclinical models of selumetinib in oncogenic KRAS-driven murine iCCA[236]. Similarly, in a phase-1 trial for the mutant IDH1 inhibitor Ivosidenib, of 73 mutant IDH1 patients only 4 showed a partial response (however data for 48 patients was censored)[237]. So far, the most successful targeted therapies are those against constitutively active FGFR2. The pan-FGFR inhibitor BGJ398 is currently in phase III trials[238] following an 82% disease control rate in phase II[239]. This demonstrates the utility of targeted therapy in certain cases and suggests that a deeper understanding of variability in response in other targeted therapies is essential for more granular stratification.

1.4 The Importance of Biological Context for Driver Function

Whether or not a particular mutation can induce transformation or contribute to oncogenic phenotypes is dependent on the cell of origin, its differentiation state, mutational landscape, and the functions, organisation and output of its signalling networks[240].

In the liver, where lineage plasticity is pronounced, the cell of origin cannot be identified in human iCCA samples. However, the cell of origin demonstrably modifies the output of oncogenic signalling. As discussed above, an example of this is the aberrant activation of the Hippo pathway. In cholangiocytes Hippo dysregulation serves to induce moderate proliferation without alteration of cell lineage but in hepatocytes induces dedifferentiation and malignancy due to expansion of the resultant progenitor pool[89]. Such dedifferentiated lesions can then acquire iCCA or HCC identity dependant on subsequent activation of different pathways [76].

Intrinsic barriers to transformation in response to different oncogenes can necessitate additional mutations, and this is also cell-type specific. For example, in the pancreas $KRAS^{G\ 2D}$ mutations are pervasive and alone can induce the acinar-to-ductular metaplasia which is the first step in pancreatic ductal adenocarcinoma[240,24] . By contrast, in hepatocytes oncogenic K/NRAS mutations result in oncogene-induced senescence (OIS)[242] which must be overcome by, for example, loss of p53[243,244]. Experimentally, this OIS can be overcome by Trp53

mutations in the mouse. In human iCCA, KRAS and NRAS show a degree of mutual exclusivity with TP53, each other, and other common drivers. This suggests that in human iCCA arising from hepatocytes there may be undiscovered drivers which allow escape from OIS. Alternatively, oncogenic RAS mutations arising in cholangiocytes may be sufficient to induce iCCA on their own. N-Ras (but not K-Ras) expression is increased in senescent cholangiocytes in PSC[245] and has been shown to induce interleukin 6 (IL6) secretion[246]. IL6 functions through the JAK/STAT pathway in normal cholangiocytes and is regulated through a feedback loop involving the transcription factor SOCS3 which is epigenetically silenced in iCCA[247,248]. This leads to inhibition of apoptosis through enhanced expression of the negative apoptosis regulator Mcl-6[247]. IL6 signalling also results in increased expression of DNMT1 causing hypermethylation which may target additional tumour suppressors[248,249]. Lastly, IL6 has been shown to act through the MAPK pathway to downregulate p21 a mediator of senescence and increase Akt pathway activation through increased production of progranulin[248,250] (a homologue of which is also secreted by *O. viverrini* to induce cholangiocyte proliferation[25 ,252]). Therefore, oncogenic RAS mutations may have different consequences in iCCA contingent on the cell of origin.

This is relevant therapeutically when considering that RAS mutational status did not predict response to MEK inhibition in the above-mentioned CCA clinical trials[234,235]. This can also be seen in other cancer types: MEK inhibition in KRAS-mutant lung and pancreatic tumours induced a cell-autonomous survival pathway and abrogated the effects of inhibition[240,253]. This phenomenon was not seen in

KRAS mutant colorectal cancer[240], in which different cells types differentially phosphorylate ERK in response to oncogenic KRAS[254]. The extent of this heterogeneity in phenotypic output was elegantly delineated in Yuan *et al.*[255] where the authors knocked down 41 KRAS effector nodes in 92 cell lines and demonstrated that each cell line had its own profile of KRAS effector dependencies. Interestingly, they then showed that despite this heterogeneity, pancreatic, lung, and colorectal cancers that all harboured oncogenic KRAS could be stratified into a limited number of groups based on differential effector engagement and this caused differential response to targeted therapies. In this model, differential effector engagement was a product of both lineage and secondary mutations. This study provides strong proof of concept that common drivers act in cell-specific manor and this is linked to drug response. The above observations reinforce the importance of delineating cell-type specific actions of drivers in the different cells of origin in iCCA[256]. Different arrangements of oncogenic signalling pathways inherent to the cell type may influence patient response to targeted therapy, making stratification based on the presence of absence of a single driver suboptimal. Identifying the cell of origin in human iCCA is not currently possible, and this prevents the integration of cell-type-specific knowledge of driver function into patient stratification frameworks. However, novel methods leveraging the mutational distribution, epigenomic and passenger landscapes of tumours to infer cell of origin may aid in this endeavour[257 259]. The application of these methods to experimental systems with known cell-of-origin as well as the whole-genome sequencing of iCCA which is noticeably lacking may help to identify molecular

signatures of originating cell type in human iCCA. integration of this knowledge with known driver mutations would aid in more granular stratification.

As well as cell-type-specific organisation of oncogenic signalling networks, such signalling networks are likely modulated by low-frequency or undiscovered drivers. The classification of cancer genes as drivers or passengers provides a theoretical framework in which to study cancer genetics, with the caveat that it is largely based on extremes. Whilst there are a defined set of powerful tumour suppressors and oncogenes, and an even larger set of genes with no biologically conceivable role in cancer, the middle-ground is likely rich in genes which influence tumour phenotypes only in particular contexts. Mutational epistasis and acquisition order influence the effect of mutations[260 265]. Epistatic interactions between driver and passenger mutations are challenging to resolve in cancer[264], but have been hypothesised in two independent mathematical models to enhance tumour fitness. In one study[266], the fitness advantage conferred on the cell by a single tumour suppressor was strongly contingent on the number of passenger mutations. By introducing epistasis into the model, the authors demonstrated that through interaction with the predominant driver, passenger mutations incrementally increased the fitness of the clone, which eventually underwent rapid expansion after a passenger threshold was reached. The authors conclude that mutations deemed to be neutral in one context cannot be regarded as neutral in all contexts i.e. in the presence of a known driver. A second study[267] demonstrated the relevance of the proposed 'middle ground' mutations which are neither known strong drivers but may not be completely neutral passengers either. In their model

they showed that deleterious passengers reduced the overall fitness of the cell, but this allowed the emergence of weak tumour suppressors which had a relatively higher fitness in that context.

The above models, whilst not based on any particular cancer type, may have relevance in iCCA. The genetic heterogeneity inherent to iCCA may facilitate the emergence of mutations which in other contexts are neutral, but through epistatic interactions with a stronger driver and/or the truly deleterious passenger repertoire may confer these mutations with weak driver status. Such weak drivers may function to modulate the oncogenic signalling networks utilised by known drivers. Elucidation of such genes and the contexts in which they gain driver status is challenging but crucial for a deeper understanding of lateral dependencies which may be targeted therapeutically in the patient population.

1.5 Contemporary Approaches for Driver Discovery from Large-Scale Sequencing Data

The proliferation of exome and genome scale sequencing in large cancer cohorts has inundated the cancer genetics community with rich datasets for analysis. A central challenge inherent in this is the filtering of signal from noise and the identification and prioritisation of biologically relevant observations. Mirroring this, there is a corresponding proliferation of computational strategies which aim to aid in this endeavour[268].

Driver prediction tools employ numerous creative approaches for the detection of driver genes or mutations. These broadly fall into three categories: functional impact prediction, recurrence, and pathway or network analysis[269]. Examples of the former acting at the mutational level include PolyPhen2[270], SIFT[27] , and MutationAssessor[272] which aim to predict deleterious mutations based on violations of evolutionarily conserved nucleotides/residues. At the gene level, tools such as ActiveDriver[273,274] identify genes with enrichment for mutations at post-translational modification residues in the protein, whilst others look for biased accumulation of functionally deleterious mutations[275] or mutational clustering in the gene[276] or protein[277]. Frequency based approaches aim to identify genes which are mutated in a cohort more often than would be expected by chance alone. The most utilised of this class of tools is the MutSigCV algorithm. This tool takes into consideration that different samples have different background mutation rates (BMR), and that different regions of the genome have different BMRs. BMR in samples may differ due to variables such as the rate of cellular proliferation, and

the integrity of the DNA repair machinery. Across a genome BMRs covary with replication timing, chromatin state, transcriptional activity, and protein-DNA binding. MutSigCV takes these considerations into account and allows the user to provide their own data to aid in accurate modelling of the BMR. This allows mutation frequencies that deviate from the BMR to be more accurately predicted[278]. The last category, pathway and network-based prediction, tests for enrichment of genes in one set (i.e. cancer mutations in a cohort) in another set (i.e. a database of genes known to be involved in a pathway or process)[279]. Other approaches utilise databases of known protein-protein interactions to construct networks *de novo* from the set of mutated genes[269,280 282].

The above examples are far from a comprehensive survey but do serve to demonstrate that different, logically sound strategies can be used to mine the extensive data for biologically relevant signals. The fact that there is no agreed upon gold standard tool for driver prediction is sometimes reported as a criticism[283]. However, the assumption that there can be a perfect method which can capture the complexity of biological systems to comprehensively detect drivers is highly unlikely, as discussed below. On the contrary, the different methods for driver prediction described above can actually be an advantage as different models and methods have complementary strengths and weaknesses. This allows the selection of a method or methods which best address the question being asked and the biological data available to answer it. This is caveated by the requirement that the results of a driver prediction analysis should ideally serve as a starting point, where predictions are rigorously validated in biological systems. This is imperative

given the importance of biological context in determining driver status and driver function, as discussed above.

There is an increasingly pervasive impression across many academic disciplines that the combination of exponentially larger datasets with more sophisticated modelling and machine learning approaches will gradually solve most problems involving complex, non-linear systems[284]. This is unlikely due to the inherent limitation that the biological and statistical models upon which prediction tools are built cannot recapitulate the complexity of biological systems which are themselves incompletely understood. Despite the impressive range and sophistication of driver prediction algorithms[269], trade-offs, biases, and inaccuracies are always present; even with machine learning, the output of such tools are still correlative an independent of mechanistic biology[284,285]. This is evident in the diversity of biologically rational strategies found in driver prediction algorithms and mirrored in the divergent results each produce. This is not a criticism of big-data-centric biology, which has resulted in unprecedented rates of progress in understanding biological systems. However, this is an acknowledgement that a deep mechanistic understanding of biological processes and complexity through classical reductionist approaches must continue to be thoroughly embraced so as to best reap the rewards of this exciting golden age of data-driven cancer genetics.

In the present study, a combination of big data analysis and *in vivo* experimental screening is used to identify novel drivers of iCCA. These are subsequently shown to display driver properties only in specific contexts.

1.6 Hypothesis

The long tail of infrequently mutated genes in iCCA contains drivers which may act alone to induce cholangiocarcinogenesis or may act with more common drivers to modify the tumour phenotype. Combining computational driver prediction and *in vivo* screening will allow the identification of such rare drivers, and the contexts they operate in.

1.7 Aims

- Define a set of predicted drivers through computational driver prediction using currently existing tumour sequencing cohorts.
- Develop a method for simultaneously screening predicted drivers *in vivo* to identify novel high-confidence drivers.
- Validate high-confidence drivers, the contexts they operate in, and their contribution to tumour phenotypes.

Chapter 2 Materials and Methods

2.1 Cloning of single sgRNAs

Single sgRNAs used in Chapter 4 were cloned into either the LentiGuide-Puro or SB-CRISPR plasmids. LentiGuide-Puro was bought from Addgene (plasmid number #52963) and SB-CRISPR was acquired from our collaborator, Professor Roland Rad (Technical University of Munich, Germany), who's lab originally generated and published the plasmid[286]. Oligos encoding sgRNAs against *loxP*, *Apc, Plxnb2, Trp53, Plk1*, *Fh1*, *Rnf31*, *Nf2* or non-targeting controls ('0007') were cloned into the backbones as follows.

sgRNA	Mouse GeCKOv2 Library ID	5'->3' Sequence
loxP	N/A	CGAAGTTATATTAAGGGTTC
0007	MGLibA_66412	ATTGTTCGACCGTCTACGGG
Apc_1	MGLibA_04599	GGATCTGTATCCAGCCGTTC
Apc_2	MGLibA_04600	GACTTACTAGAGCGTCTTAA
Apc_3	MGLibA_04601	AGATCCTTCCCGACTTCCGT
Trp53_1	MGLibA_56033	AGTGAAGCCCTCCGAGTGTC
Trp53_2	MGLibA_56034	AACAGATCGTCCATGCAGTG
Trp53_3	MGLibA_56035	TGAGGGCTTACCATCACCAT
Fh1_1	MGLibA_18403	AATTGGGCGAACTCACACGC
Fh1_2	MGLibA_18404	CGTGTAGAGTTCGACACCTT
Fh1_3	MGLibA_18405	AAAATCCAAAGAGTTTGCGC
Plxnb2_1	MGLibA_41608	CTCAGATGGCCGGATCCTTA
Plxnb2_2	MGLibA_41609	TACTCTGCAGAAGTGCCGTC
Plxnb2_3	MGLibA_41610	GGAGTCACGACACTGAGCGC
Nf2_1	MGLibA_33838	GATCCGCACCGTGAATGTCT
Nf2_2	MGLibA_33839	GAAGCTCATGCGAGAAGCGA
Nf2_3	MGLibA_33840	CTTGGCGTCATATGCTGTCC
Rnf31_1	MGLibA_45954	CATACAACCGTAGTACATCC
Rnf31_2	MGLibA_45955	AGGGTGGCCGGGATGTACTA
Rnf31_3	MGLibA_45956	TAACCCCGTCTTTCGCAGCA
Plk2_1	MGLibA_41524	CAGAAGTCCGATACTACCTC
Plk2_2	MGLibA_41525	AAAACTGCACGACATGCTTA
Plk2_3	MGLibA_41526	ACGAACAAGAAATCTTGCAC

Table 2.3.1 sgRNA sequences for single-guide cloning

The LentiGuide-Puro or SB-CRIPSR backbone was digested with Esp3I (New England Biolabs; R0734S) or BbsI (New England Biolabs; R3539S), respectively, to produce staggered ends, with simultaneous dephosphorylation using FastAP Thermosensitive Alkaline Phosphatase (ThermoFisher Scientific; EF0654) to prevent backbone re-closure during ligation. Plasmid digestion products were size separated by gel electrophoresis in 2% agarose gel for 45-60 minutes at 120V (constant), and extracted/purified (Zymoclean Gel Recovery Kit, Zymo Research; D4008). sgRNA oligos were annealed and simultaneously phosphorylated with T4 Polynucleotide Kinase (New England Biolabs; M0201S) for 30 minutes at 37^{o}c. Ligation of plasmid backbone and insert was carried out at 16^{o}c overnight with T4 ligase (New England Biosciences; M0202S). Ligation reactions were heat-inactivated and transformed into Stbl3 cells as per the manufacturer's instructions (One Shot Stbl3 Chemically Competent *E.coli*; ThermoFisher Scientific; C737303) and grown on Luria Broth agar at 37^{o}c overnight. Sanger sequencing of colonies (MRC Human Genetics Unit Technical Services) identified those with correctly ligated plasmids which were subsequently picked, cultured and extracted as maxi-prep's as per the manufacturer's instructions (QIAfilter Plasmid Maxi Kit, QIAGEN; #12263). Plasmid yield was assessed on the NanoDrop ND-1000 Spectrophotometer (ThermoFisher Scientific).

2.2 CRISPR/Cas9 Screening Library Generation

Oligonucleotides encoding sgRNAs targeting the set of predicted drivers were designed using spacer sequences from the mouse GeCKo V2 library[287] with the

Python script supplied in the corresponding paper[288] (see Appendix 1 for all sgRNA sequences). Library-specific PCR retrieval arms were derived from the protocol for parallel oligonucleotide retrieval published by the Luo lab[289]. Table 2.3.1 and 2.3.2 show the general schematic for both libraries, and the primers for specific library amplification, respectively.

Lib	Forward Primer-Binding Arm	5' Esp3I cut site	5'->3' Sequence	3' Esp3I cut site	Reverse Primer-Binding Arm
A	GCACAGTTTCACGATCGCTTTAGC	CGTCTCACACCG	N(20)	GTTTTGAGACG	CTGTCGTGGAATCGCTAAAGCAAG
B	ACGTGGATGCTAGTGCGGAGTTTA	CGTCTCACACCG	N(20)	GTTTTGAGACG	TAGCGTATGACGGCGTTCTGCTAT

Table 2.3.1 Oligo schematic for library pool synthesis

Library	Forward Primer	Reverse Primer
A	GCACAGTTTCACGATCGCTTTAGC	CTTGCTTTAGCGATTCCACGACAG
B	ACGTGGATGCTAGTGCGGAGTTTA	ATAGCAGAACGCCGTCATACGCTA

Table 2.3.2 Primers for amplifying individual libraries from starter pool

Complete sgRNA oligos for all target genes and control sequences were custom synthesised by Twist Biosciences. Library-A and library-B oligos were amplified from the custom oligo pool separately with 22 cycles on a thermocycler with PCR reagents (NEBNext HF PCR MasterMix, New England Biosciences; M0541S). Products were purified using the QIAquick Nucleotide Removal Kit (QIAGEN; #28304). Purified oligos were then digested with Esp3I and phosphorylated as per the single sgRNA cloning procedure above. Digested inserts were purified from the reaction by isopropanol precipitation. For each library, this process was repeated for 3 technical replicates, which were then pooled together per-library before ligation into the SB-CRISPR plasmid backbone. This was to ensure that all oligos were present in excess to limit biasing the library downstream. Ligation of inserts into digested SB-CRISPR was carried out in 3 technical replicates which were ligated

overnight at 16°c with T4 ligase as above and then pooled together. Transformation of ligated products into Stbl3 was carried out for 10 technical replicates, once again to ensure unbiased and complete library representation. Transformants were grown overnight at 37°c on Luria Broth agar in 500cm^2 square plates (ThermoScientific Nunc Square Culture Dish; #UY-01929-00). All colonies were scraped into flasks and cultured for a further 6 hours in LB at 37°c before being pelleted and maxi-prep'd. Resultant libraries underwent amplicon sequencing to determine the representation and evenness in sgRNA sequences. sgRNA inserts were PCR amplified with primers targeting the sequence before and after the insert site. This step also served at add MiSeq (Illumina) flow-cell adaptors, variable length random sequences to increase sequencing library complexity, and index sequences for multiplexed sequencing. Sequencing quality control, sgRNA quantification, read coverage and Gini coefficients were generated from unprocessed FASTQ files using the MAGeCK-VISPR software[290].

2.3 Lentivirus Generation

Lentiviruses for delivery of LentiCas9-Blast (Addgene #52962), dCas9, or LentiGuide-Puro-sgloxP and -sg0007 (non-targeting control) to biliary organoids were made as follows. Lentiviral vector mix was made by combining 7.5μg of PAX2 helper vector (Addgene #12260), 2.5g of pCMV-VSV-G envelope plasmid (Addgene #8454), and 10μg of lentiviral vector in 500μl Advanced Dulbecco's Modified Eagle Medium (ThermoFisher Scientific; #12491015). This was added to a mix of 425μl DMEM and 75μl 1mg/ml polyethylenimine (PEI linear, 25K molecular weight,

Polysciences Inc.; 23966-1). This combined mixture was added to HEK293 cells which were cultured overnight. The following day media were collected and lentivirus was concentrated using the Lenti-X lentiviral concentrator (Takara; #631232) and titrated to determine concentration using the Lenti-X qRT-PCR Lentivirus Titration Kit (Takara; #631235).

2.4 DNA and RNA extraction from tissue

Both DNA and RNA extraction protocols begin with 50-100mg of snap frozen tissue. DNA was extracted from tissue using the DNeasy Blood and Tissue Kit (QIAGEN; #69504) as per the manufacturer's instructions. RNA was extracted by placing the tissue into 500μl TRIzol RNA Isolation Reagent (Invitrogen; #15596026), which was then homogenized for 3-10 minutes with a steel bead in the Qiagen TissueLyser LT (QIAGEN; #69980) depending on the toughness of the tissue (tumours were very stromal and solid compared to control liver, for example). To this, 100μl chloroform was added and the sample was centrifuged at 12,000 RPM at 4°c for 15 minutes. The clear upper fraction was decanted into a fresh tube and an equal volume of isopropanol was added. This was then processed using the RNeasy Mini Kit (QIAGEN; #74104) as per the manufacturer's instructions from step 5 of the protocol onwards. For downstream sequencing applications DNA and RNA quality (RIN score) was quantified using the Agilent 2100 Bioanalyzer with either the DNA 1000 chip, or RNA 6000 chip (MRC Human Genetics Unit Technical Services). A minimum RIN threshold of 8 was used for RNA-seq.

2.5 Real-Time Quantitative PCR

1μg of RNA was reverse transcribed into cDNA using the QuantiTect Reverse Transcription Kit (QIAGEN; #205311). RT-qPCR was subsequently performed using the LightCycler 480 SYBR Green-1 Master Mix (Roche; #04707516001) on the LightCycler 480 machine (Roche). Primers against Cas9 were designed and validated using NCBI Primer Blast by Nerea Galdona Yeregui, a Master's student I co-supervised in the lab. Primers targeting housekeeping genes peptidylprolyl isomerase A (PPIA) and glyceraldehyde 3-phosphate dehydrogenase (GAPDH) were from the QIAGEN QuanitiTect Primer Assay Library. Each sample was assayed in triplicate for each gene to be quantified, and DNase/RNase-free water was used as a control. Quantification and calculation of expression fold change was carried out using the double delta Ct method[29] .

Gene	Forward Primer	Reverse Primer
Cas9	AAAGACCGAGGTGCAGACAG	ACCACCAGCACAGAATAGGC
PPIA	Mm_Ppia_1_SG QuantiTect Primer Assay (#QT00247709)	
GAPDH	Mm_Gapdh_1_SG QuantiTect Primer Assay (#QT00199388)	

Table 2.6.1 sequences and assays for qPCR

2.6 Exome Sequencing of Mouse Tumours

DNA was extracted from tumours and one normal liver as described above and sent for exome sequencing at the Edinburgh Wellcome Trust Clinical Research Facility.

2.7 Transcriptome Sequencing of Mouse Tumours

RNA was extracted from tumours and one normal liver as described above. Total RNA sequencing with ribonucleotide depletion, but without PolyA enrichment, was carried out by the Wellcome Trust Clinical Research Facility. This modality was chosen over polyA enrichment so as to capture non-coding RNA species.

2.8 Generation of $R26^{mT/mG}$ (d)Cas9-expressing Bile Duct Organoid lines and Lentiviral Transduction.

Livers were perfused with phosphate buffered saline and dissected out of $R26^{mT/mG}$ mice. The livers were minced into ~1mm^2 chunks which were then digested in digestion media (100ml organoid base media with 12.5mg collagenase XI (Sigma; #C9407) and 12.5mg dispase II (Gibco; #17105-041)) for 4-6 hours with frequent agitation. Bile duct fragments were picked and washed in PBS before being suspended in Growth Factor-Reduced Matrigel (Corning; #356252) and plated in 12-well sterile culture plates (Cellstar, Greiner Bio-One; #665-180) and immersed in organoid media. After several days, organoids were passaged and subsequently expanded for 2-3 further passages before being scraped from the plates and incubated in ice cold versene for 30 minutes at 4^{o}c. This served to dissolve the Matrigel. The organoids were washed 3 times in ice cold PBS and fragmented with a pipette. A layer of Matrigel was spread to cover the bottom of the wells in a pre-warmed 6-well culture plate and left to polymerise. 500μl of organoid growth factor media was mixed with either 0.2μl, 0.5μl or 1μl of Cas9 or dCas9 lentivirus and 2.5μl TransDux transduction reagent (System Biosciences; #LV850A-1). Organoid

fragments were pelleted and the supernatant was removed before pellets were resuspended in the lentiviral mixture and plated on the Matrigel lawn. After overnight incubation the lentivirus/media mix was removed and the surface was washed with PBS to remove dead cells before a layer of Matrigel was overlain to allow the organoid fragments to re-form organoids. Passaging with puromycin selection produced a Cas9 or dCas9-expressing organoid line, as confirmed by qPCR. Transduction of organoids with sgRNAs for *mT/mG* locus editing was performed as above with LentiGuide-Puro-sgloxP/-sg0007 lentiviral particles.

Component	Volume
Advanced DMEM/F12	500ml
Penicillin/Streptomycin	5ml
HEPES (1M)	5ml
Glutamax (100x)	5ml
Supplement	**Concentration**
R-Spondin1	4ul/ml
hEGF	0.1ul/ml
hFGF-10	0.1ul/ml
HGF	0.1ul/ml
Gastrin	0.1ul/ml
Nicotinamide	10ul/ml
N-acetylcytosine	2.5ul/ml
B27 (without Vitamin A)	20ul/ml
ROCK inhibitor (Y-27632)	2ul/ml
Forskolin	1ul/ml
TGFβ inhibitor (A83-01)	0.5ul/ml
Gsk3β inhibitor (Chir99021)	0.3ul/ml
Recombinant Wnt3a	1ul/ml

Table 2.9.1 Organoid Growth Media

2.9 Mouse Hydrodynamic Tail Vein Injections

Animal work was carried out primarily by Dr Luke Boulter as I do not hold an animal licence. 20ug NICD plasmids, $RAS^{G\ 2}$ plasmids, and SB-CRISPR-library/single-sgRNA plasmids were injected per mouse. 6ug SB13 transposase plasmid was injected per mouse. Plasmids were delivered over the course of 7-10 seconds through the tail vein.

2.10 Tissue Collection and Processing

After culling, livers were perfused with PBS through the inferior vena cava with drainage through the hepatic portal vein, and subsequently dissected out of mice. Livers were photographed. When present, tumours were dissected out of livers and, if large enough, were cut into 3 equal sections for histology, DNA, and RNA extraction. DNA and RNA extraction were prioritised over histology when tumours were small. Tissue for DNA was snap frozen in 2-methylbutane chilled on dry ice and subsequently stored at -80^{o}c. Tissue for RNA was perfused overnight in RNAlater (ThermoFisher Scientific; #AM7020) at 4^{o}c before being stored at -80^{o}c. Tissue for histology was fixed in 4% formaldehyde (VWR Chemicals; #9713.5000) for 24 hours before being transferred to 70% ethanol for a minimum of 24 hours. Tissue was then perfused (Tissue-Tek VIP 5 Jr.) and mounted in paraffin (Tissue-Tek III Paraffin WAX; #4509 and Tissue-Tek Cassette System) and sectioned as required.

2.11 Immunohistochemistry

Sections were dewaxed in 100% xylene for 5x2 minutes and rehydrated through a series of solutions of decreasing ethanol concentrations (100%, 75%, 50%, 25%) for 5 minutes each before being placed in water for 15-20 minutes. Following antigen retrieval samples were cooled in gently running cold water for 10 minutes and washed in 3x5 minute PBS washes. Samples were then incubated at room temperature for 15 minutes in 3% hydrogen peroxide and washed for 3x5 minutes in PBS before being mounted on coverplates (Thermo Scientific Shandon Plastic Coverplates; #72110017) and placed in Sequenza racks (VWR; #10129-584). Endogenous avidin binding sites, biotin, and biotin receptors were blocked with an avadin/biotin blocking kit sequentially for 15 minutes at room temperature (Avadin/Biotin Blocking Kit, Vector Laboratories; SP-2001). Non-specific staining was decrease by protein blocking for 30 minutes at room temperature with a pan-species protein block (Agilent-Dako; X090930-2). Primary antibodies were diluted in antibody diluent (Antibody diluent, Life Technologies, ThermoFisher Scientific; #00-3218) and incubated overnight. Following 3x5 minute PBS washes, biotinylated secondary antibody was diluted in antibody diluent and incubated with the sections for 30 minutes. Horseradish peroxidase was conjugated to the secondary using the VECTASTAIN ABC HRP system (Vector Laboratories; #PK-4010) and DAB detection was performed by incubating the slides for 3-5 minutes with DAB chromogen (DAB Quanto Chromogen and Substrate, ThermoFisher Scientific; TA-060-QHDX). Slides were counterstained with Harris haematoxylin and developed with Scott's tap water before being dehydrated through increasing alcohols (as

above, in reverse) and cleared for 10 minutes in xylene. Slides were mounted with DPX (CellPath; SEA-1304-00A).

Antigen	Antigen Retrieval	Primary Antibody	Secondary Antibody
Pan-Cytokeratin	Sodium Citrate	DAKO (Z0622); 1:300	Biotinylated anti-rabbit
Cytokeratin 19	Sodium Citrate	DSBH (AB2133570); 1:200	Biotinylated anti-rat
Collagen I	Sodium Citrate	Southern Biotech (1310-01); 1:1000	Biotinylated anti-goat
Collagen VI	Sodium Citrate	Merck (HPA045239-100UL); 1:100	Biotinylated anti-rabbit
Pan-Fibronectin	Sodium Citrate	AbCam (Ab2413); 1:200	Biotinylated anti-rabbit
α-Smooth Muscle Actin	Sodium Citrate	Sigma (A5228); 1:2000	Biotinylated anti-mouse
F4/80	Proteinase K	AbCam (Ab6640); 1:400	Biotinylated anti-rat
Foxp3	Sodium Citrate	Invitrogen (FJK-16s); 1:500	Biotinylated anti-rat
MPO	Sodium Citrate	AbCam (Ab9535); 1:100	Biotinylated anti-rabbit
GFP	Tris-EDTA	Santa Cruz (sc-8334); 1:500	Biotinylated anti-chicken

Table 2.11.1 Antibodies and their antigen retrieval conditions.

Computational Methods

Unless otherwise stated, analyses were conducted using the Eddie Mark 3 High-Performance Compute Cluster running Scientific Linux 7 (‘Eddie3’) at the University of Edinburgh or a MacBook Pro with 2.2 GHz Intel Core i7 processor and 16 GB memory running macOS versions 10.xx.xx (‘desktop’).

2.12 Statistical Analysis

All statistical analyses were conducted using either R versions 3.0 3.6, or GraphPad Prism 7 on a desktop.

2.13 Published Exome-seq Data Acquisition

Exome-seq FASTQ files from Chan-on *et al.* [68] were downloaded from the European Nucleotide Archive with accession PRJEB4445. Exome-seq FASTQ files from Sia *et al.*[222] were downloaded from the Gene Expression Omnibus Database with accession GSE63420. TCGA BAM files were downloaded from the Genomic Data Commons after receiving access to individual patient BAM files.

2.14 Mutation Calling

For the TCGA, Chan-On, and Sia datasets, mutations were called as follows. Input FASTQ files were aligned to the Hg19 reference genome using Burrows-Wheeler Aligner[292] version 0.7.15. PCR and optical duplicates were marked using Picard MarkDuplicates version 2.5.0, base quality score recalibration was carried out using BaseRecalibrator and local indel realignment was performed with IndelRealigner (both from the Genome Analysis Toolkit[293] version 3.6 (GATK3)). Ensemble variant calling was performed with Mutect[278] (1.1.5), Mutect2 (2.1), Freebayes[294] (1.0.2.29), Vardict-java[295] (1.4.6), and Varscan[296] (2.4.2) and all passing somatic variants which were identified by 2 or more algorithms were taken forward. The above analysis was batched and implemented within the Blue Collar Bioinformatics pipeline (bcbio-nextgen 1.0.0a0-af4730e). TCGA data were input as reads which were aligned to the Hg38 reference so, following mutation calling, variants were re-mapped using Picard LiftoverVCF (2.5.0) using the Hg38toHg19 chain file provided by UCSC. Variant annotation was performed with the Ensembl Variant Effect Predictor[297] (v88) using the '--pick', '--tab', and '--symbol' options.

2.15 Driver Prediction with IntOgen-Mutations

IntOgen-mutations[275,276,298,299] was installed on Eddie3 by creating a virtual environment with Anaconda (4.3.1). The Anaconda environment was created with the following packages to run IntOgen-mutations:

Package	Version
ago	0.0.9
bgconfig	0.5.0
bgdata	1.0.0
certifi	2017.4.17
configobj	5.0.6
cycler	0.10.0
decorator	4.0.11
drmaa	0.7.7
homura	0.1.5
humanize	0.5.1
intogen	3.0.8
libgfortran	1
matplotlib	2.0.0
networkx	1.11
numpy	1.8.2
numpy	1.12.1
oncodriveclust	1.0.0
oncodrivefm	1.0.3
openssl	1.0.2k
pandas	0.17.0
patsy	0.4.1
pip	9.0.1
pycurl	7.43.0
pyliftover	0.3
pyparsing	2.2.0
python	3.4.5
python-dateutil	2.6.0
pytz	2017.2
readline	6.2
ruffus	2.6.3
scipy	0.14.0
setuptools	27.2.0
six	1.10.0

Table 2.15.1 Packages used to create the Intogen Anaconda virtual environment.

Additionally, MutSigCV[300] (1.3.01) was installed with its Matlab Runtime/Compiler dependency (Matlab MCR 9.0.1v901) and integrated into the IntOgen-mutations pipeline as per the IntOgen documentation (bitbucket.org/intogen/intogen-pipeline/scr/master/). IntOgen was run on each published/TCGA cohort individually, and then all combined. Significance thresholds for OncoDriveFM and OncoDriveCLUST were q=0.05, and q=0.1 for MutSigCV as per the default recommendations. The minimum frequency of occurrence for a gene to be considered for analysis was n=2 for OncoDriveFM and n=5 for OncoDriveCLUST.

2.16 Assigning Mutational Signatures

The DeconstructSigs[30] package was used in R to infer mutational processes active in samples which had a number of mutations out-with the 3rd quartile for that cohort's mutation distribution. Samples in this set with predominant mutational signatures[302] linked to hypermutator processes were removed.

2.17 Functional Interaction Inference

Cytoscape[28] (3.7.1) with the ReactomeFIVis[303] app was used to build a network of known and inferred functional interactions between the set of predicted drivers so as to explore the relationship between them in a cellular context. This network was clustered into modules by connection density and each module was annotated with pathway enrichments. This was overlain with cancer drugs using the "Fetch Cancer Drugs" functionality.

2.18 CRISPR/Cas9-editing validation and Structural Variant Analysis

DNA was extracted from FASTQ files from exome-sequencing of tumours arising from the RAS$^{G\ 2}$-library screens were aligned to the FVB mouse reference genome; subsequently, indels within 50bp upstream or downstream of sgRNA target sites called. This was done by the IGMM Bioinformatics Core (Dr Alison Meynert). The following was done by me.

To determine if indels were likely due to Cas9 activity or were spurious mutations, the interval of each indel was observed on the FVB genome using Integrative Genomics Viewer[304], and overlain with sgRNA library-A/B binding sites. Indels with start or end sites out-with sgRNA targets were removed, although this was only the case for 5 indels across all samples.

To determine if editing had induced structural rearrangements, structural variants (SVs) were called using Delly2[305] (v0.7.9) as follows: all reads which had a minimum mapping quality of 20 were used to call germline and somatic SVs, SVs with a PASS flag for somatic status, had a minimum variant allele fraction of 0.15, had zero read support in the normal, and had precisely mapped breakpoints were then visually analysed on IGV to determine if they overlapped with sgRNA target sites.

2.19 Transcriptome Sequencing Analysis

Analysis of RNA-seq data from tumours was done by the IGMM Bioinformatics Core (Dr Graeme Grimes) to produce PCA plots and differentially expressed gene sets. Volcano plots were made by me.

2.20 Functional Enrichment Analysis, Clustering, and Heatmaps

Driver gene and differentially expressed gene functional enrichment was done using g:Profiler[306]. For differentially expressed genes the set was analysed as an ordered query ordered by q-value. Enrichment for Gene Ontology[279,307] Molecular Function, Biological Process, and Cellular Component terms was carried out, along with KEGG[308] and Reactome[282] pathway enrichment, enrichment for transcription factor binding sites using TRANSFAC[309] and enrichment for miRNA interactions using miRTarBase[310].

Dendrograms and heatmaps were generated in R using the Pheatmaps package (https://cran.r-project.org/web/packages/pheatmap/index.html) with base R hclust 'complete' clustering on row and column Euclidean distance matrices.

2.21 Figures

Individual figures were largely made using the R packages ggplot2[313], Pheatmap, and EnhancedVolcano (DOI:10.18129/B9.bioc.EnhancedVolcano), or GraphPad Prism 7 or Affinity Designer. All figure layouts were assembled in Affinity Designer.

Where possible, the colour-blind optimised Viridis palette (https://cran.r-project.org/web/packages/viridis/index.html) was used.

Chapter 3 Re-analysis of existing sequencing data uncovers a novel set of candidate driver genes

3.1 Introduction

The proliferation of exome sequencing (exome-seq) studies in iCCA has provided a rich source of data which can be mined for biologically relevant signals. In this chapter the set of published exome-seq resources are curated and harmonised into a unified set which is subsequently used for driver prediction. A literature search identified three driver prediction tools which had complementary strengths and weaknesses and were well suited for the data and question being asked. Using these on the combined input set identified common, known iCCA drivers, as well as a large set which are found at low frequency in the patient population.

3.2 Hypothesis

Candidate low-frequency iCCA driver genes can be identified from published exome-sequencing cohorts using complementary driver prediction methods.

3.3 Aims

- Identify and characterise suitable sources of iCCA exome-sequencing data from published studies and data repositories.
- Identify and apply suitable methods for identifying candidate driver mutations which do not rely on mutation frequency alone.

3.4 Existing data sources capture key epidemiological trends

Published exome-seq studies have used cohorts of various sizes (range:8-239) to explore questions in biliary tract cancer genetics [64, 65, 68,209,222]. To best utilise the existing data, datasets from the sources listed in table 3.4.1 were reviewed to characterise the demographic and clinical characteristics of the samples. This made use of both individual published cohorts and data generated as part of the agglomerative project The Cancer Genome Atlas (TCGA) and identified a set suitable for inclusion in subsequent driver prediction analysis.

Study	PubMed ID	Sample Size	Subtype	Samples After Filtering
Nakamura	26258846	239	iCCA, eCCA, GBC	125
Zou	25526346	103	iCCA	96
TCGA	N/A*	51	iCCA, eCCA, GBC	39
Chan-on	24185513	15	iCCA, eCCA	9
Sia	25608663	8	iCCA	8
*see Farshidfar *et al.* for analysis on a TCGA subset		416		277

Table 3.4.1 The five sources of exome-seq data used in the present study. TCGA data are accessed from the repository. Filtering for correct subtype, fluke, HBV/HCV integration, and PSC status, and hypermutation status reduces the total input to 277.

The Nakamura [65], TCGA, and Chan-On [68] data contained a mix of iCCA, extrahepatic cholangiocarcinoma (eCCA) and gall bladder carcinoma (GCA), and samples of the latter subtypes were excluded. As discussed in the Introduction, the risk factors associated with iCCA are varied and incompletely understood. Because the link between fluke infection and iCCA development is the best characterised out of known risk factors, the decision was made to remove samples with reported fluke involvement. This was deemed appropriate as the aim of this study was the

identification of novel drivers in ‘spontaneous’ disease: those cases without a clear and established predisposing risk factor. As such, a single sample with reported PSC was also removed as this is another well-defined risk factor.

Samples with less strongly associated risk factors such as alcohol consumption, tabaco exposure, and anti-HBV/HCV antibody reactivity were not excluded except in one case where HCV tumour genome integration was confirmed in the respective study.

Generally, across studies there were incomplete clinical and demographic metadata associated with each patient, and the types of metadata which were reported varied between sources (Fig 3.1). In some instances, basic variables such as sex, age, and ethnicity were not reported. This may be the result of stipulations made by ethics boards regarding patient anonymity and/or due to the irrelevance of reporting such clinical or demographic data to the goals of the respective studies.

The combined sample set had a preponderance towards male patients without any reported known risk factors. The age-at-diagnosis frequency distribution is centred around a mean whole-number age of 57 years (range: 29-86), which differs significantly from the peak incidence of the 7^{h} decade reported in recent epidemiological studies[26] ($P<0.0001$, one-sample t-test). HCV and HBV are present in a proportion of the patients as would be expected due to their association with iCCA: a greater percentage of samples have HBV than HCV and this likely reflects the endemic status of this virus in parts of Asia [26], where most of the studies were carried out. Smoking and alcohol consumption are not recognised as strong risk

factors and this is reflected in their low incidence here: 6.12% and 3.60%, respectively. Overall, the combination of samples currently available from already published studies constitutes a set that closely captures key epidemiological trends regarding age, ethnicity, sex, and risk factors.

One key consideration which should be highlighted is that in contrast to other cancer types such as oesophageal, colorectal, leukaemia and lymphoma iCCA and liver malignancies generally are not easily accessible without surgical procedures. As such, all of the samples included here were collected in their respective studies as a by-product of surgery with curative intent. No samples were reported to have been taken post-mortem or otherwise out-with the context of surgical resection. A consequence of this is that all current sequencing data for iCCA is representative of only the small minority of iCCA patients who meet the criteria for resection. Subsequently they may be unintentionally enriched for less aggressive tumours without extensive vascular or lymph node involvement, or that were detected early (a possible explanation for the lower than expected average age in the cohort). This is a challenging problem to overcome and has yet to be widely explored in the iCCA field, or the cancer research field more generally.

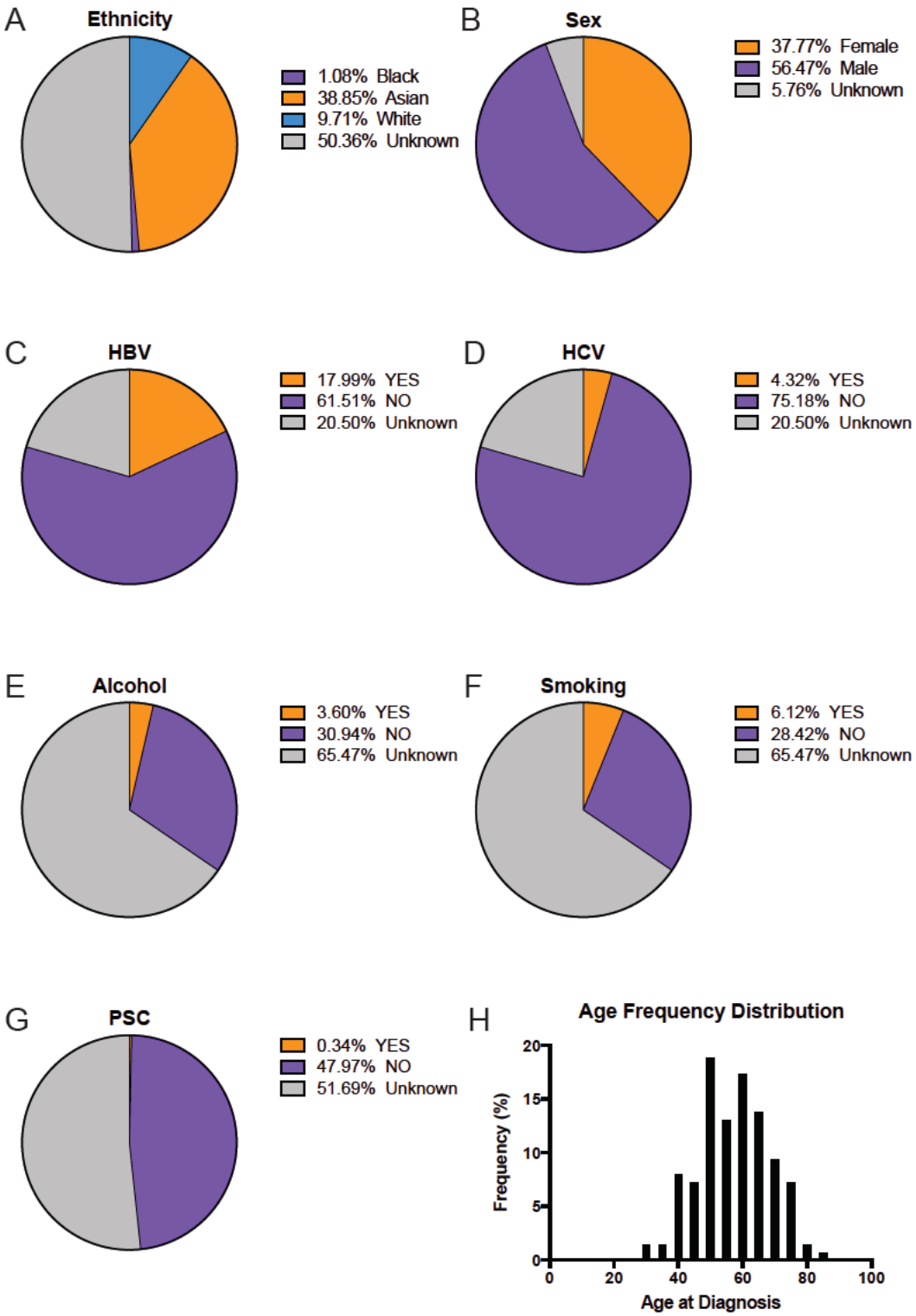

A
Ethnicity
1.08% Black
38.85% Asian
9.71% White
50.36% Unknown
B
Sex
37.77% Female
56.47% Male
5.76% Unknown
C
HBV
17.99% YES
61.51% NO
20.50% Unknown
D
HCV
4.32% YES
75.18% NO
20.50% Unknown
E
Alcohol
3.60% YES
30.94% NO
65.47% Unknown
F
Smoking
6.12% YES
28.42% NO
65.47% Unknown
G
PSC
0.34% YES
47.97% NO
51.69% Unknown
H
Age Frequency Distribution
Frequency (%)
0
5
10
15
20
0
20
40
60
80
100
Age at Diagnosis

Figure 3.4.1 Demographic and clinical attributes of the combined cohort: A-G shows the proportion of each clinical attribute in the combined cohort; notably there is a lot of missing data which are not reported in the respective publications. H, distribution of ages of patients in the pooled cohort, the average is centred around 57 years younger than the average age at diagnosis in larger epidemiological studies.

3.5 Data harmonisation produces a unified mutation set for driver prediction

Sequencing data from each source were available in various formats and at various stages of processing. Ideally, data would be acquired as raw sequencing reads with little to no processing. This would aid in data harmonisation between the different sources by allow maximum control over bioinformatic processing thus limiting factors affecting the number of mutations called to library preparation and sequencing variables (DNA quality, exome capture probes used, depth of coverage, sequencing platform, etc.). In reality however, this was not possible as raw reads were not readily available for all samples. Sequencing reads from the Zou [64] dataset were accessible on the Short Reads Archive (SRA) however the correspondence between files and biological samples, and between tumour and matching normal reads was not self-evident, nor was it explained in the publication or its supplementary data. Repeated attempts to gain clarification from the authors through personal communication proved fruitless. Sequencing reads from the Nakamura dataset were available through the International Cancer Genome Consortium (ICGC), however, at the time of the present analysis the complete set of raw reads corresponding to the iCCA samples reported in the paper were not accessible (at the time of writing the complete data are now available). To expedite the analysis in light of the above issues the decision was made to derive mutations from the publications' supplementary data.

The mutations reported for the above studies had undergone stringent filtering and only high-confidence single nucleotide variants (SNVs) or small insertions or

deletions (indels) were reported. For each set, single nucleotide polymorphisms found in dbSNP were removed if present and the remaining mutations were converted into a format suitable for downstream analysis using Python programming.

For the Sia[222], Chan-On, and TCGA samples, data were available as either unaligned sequencing reads (FASTA format) or aligned reads (BAM format) and in those cases mutations were called using an in-house pipeline (Materials and Methods).

Study	Data Format	Mean muts/sample	Min-Max
Nakamura	Supplementary	330	10 - 6119
Zou	Supplementary	70	1 - 1268
TCGA	BAM	977	416 - 7868
Chan-On	FASTA	400	272 - 549
Sia	FASTA	381	166 - 670

Table 3.5.1 Data format and statistics of mutations in exome-seq sources

The average number of mutations differed between sources and the reason for this is not entirely resolvable but is likely explained in part by the in-house mutation calling strategy. The in-house analysis employed a pipeline with an ensemble mutation calling step wherein mutations were called using 5 different algorithms. Mutations which were independently called by 2 or more methods whilst passing their internal filters for somatic mutations were carried forward for driver prediction. In contrast with the Nakamura and Zou mutations, there was no hard filtering to minimise false-positives after initial calling. The goal of this computational analysis was not to derive a list of very high-confidence mutations, but to provide input for a pipeline which ultimately undergoes experimental validation. The expectation with this strategy is that false-positive calls present at

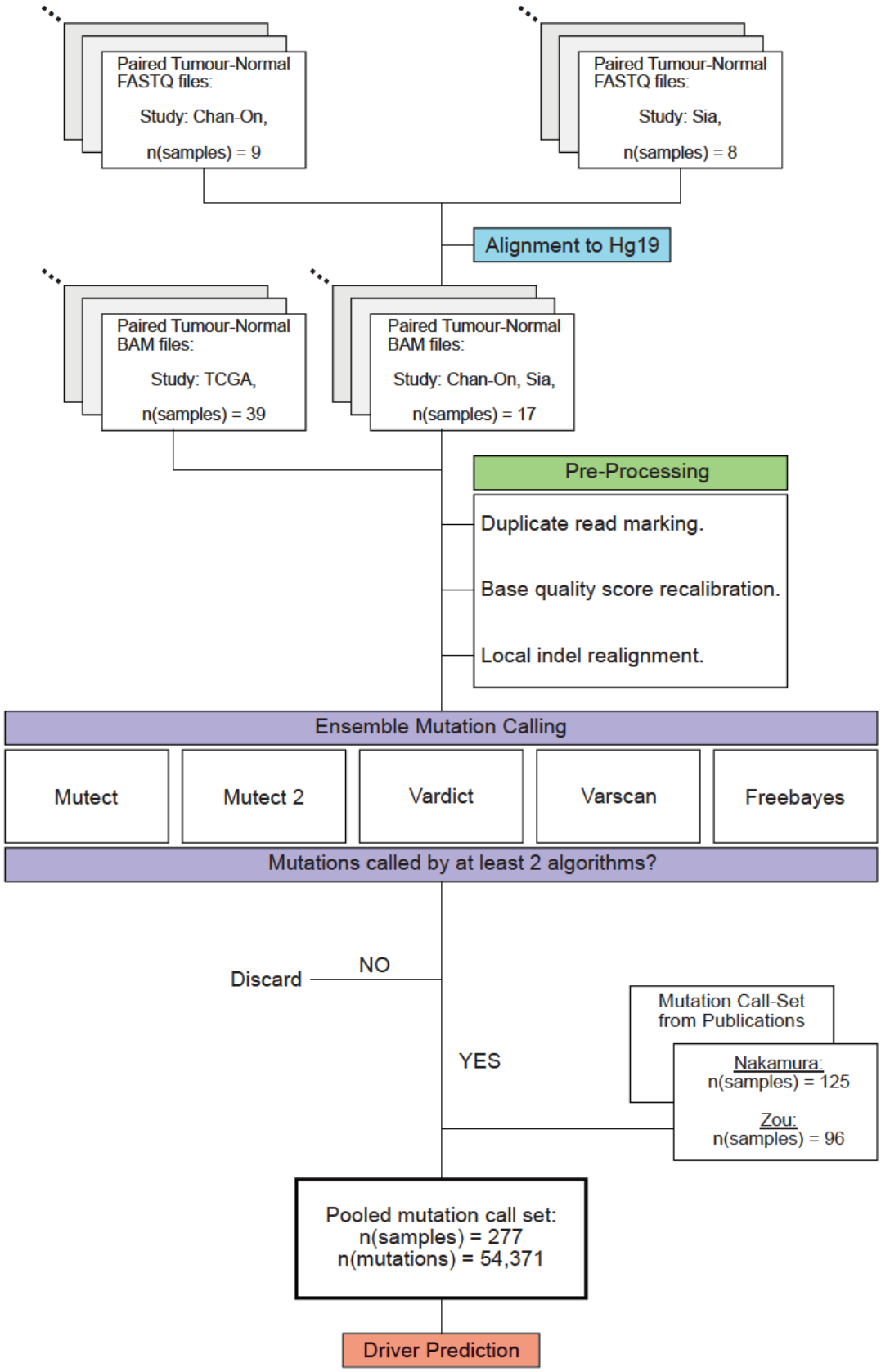

Paired Tumour-Normal FASTQ files:
Study: Chan-On,
n(samples) = 9
Paired Tumour-Normal FASTQ files:
Study: Sia,
n(samples) = 8
Alignment to Hg19
Paired Tumour-Normal BAM files:
Study: TCGA,
n(samples) = 39
Paired Tumour-Normal BAM files:
Study: Chan-On, Sia,
n(samples) = 17
Pre-Processing
Duplicate read marking.
Base quality score recalibration.
Local indel realignment.
Ensemble Mutation Calling
Mutect
Mutect 2
Vardict
Varscan
Freebayes
Mutations called by at least 2 algorithms?
Discard
NO
YES
Mutation Call-Set from Publications
Nakamura:
n(samples) = 125
Zou:
n(samples) = 96
Pooled mutation call set:
n(samples) = 277
n(mutations) = 54,371
Driver Prediction

Figure 3.5.1 Data processing pipeline. The data were batch processed such that FASTQ files were aligned to Hg19 reference genome before undergoing pre-processing to correct for artefacts and systemic effects. Mutations were then called by 5 algorithms and those mutations which were identified by 2 or more algorithms were taken forward for driver prediction. See Materials and Methods for details.

this stage and which may be propagated through the driver prediction steps will be lost during functional screening if they are not relevant to tumour initiation or progression. The advantage of this approach is that mutations which recurrently arise later in tumour evolution and are present subclonally will be retained whereas they may normally be lost due to hard filtering on read depth or variant allele frequency. The identification of subclonal mutations relies on very high depth sequencing[3 2] which has not been reported in iCCA, however this more relaxed approach to mutation filtering may retain genuine subclonal drivers lacking traditionally adequate read support and allow them to emerge during experimental screening.

Whilst this may account for the discrepancy between the in-house sets and the supplemental sets it does not explain the large difference in average between the Sia/Chan-On sets and the TCGA set. These 3 sets were processed with the same pipeline with the exception of an additional step for TCGA samples. The original aligned reads from TCGA were aligned to the GRCh38 reference genome but mutation coordinates were converted to the GRCh37 release to aid in data harmonisation with the other sources and compatibility with downstream analyses. This step produced no error or unexpected outcome, and visual inspection of 50 randomly selected mutations using UCSC Genome Browser confirmed that this had been successful. The only other variable which may explain the discrepancy is read depth and coverage being substantially greater in TCGA samples compared to the Sia and Chan-On samples. Explicit sequencing metrics were not reported in the latter studies and given their small size it may be

reasonable to assume that sequencing depth was lower than in TCGA samples although this cannot be confirmed.

The above bioinformatics workflow harmonised data for input to subsequent analyses by calling mutations using a common framework where possible, and by recording mutations for each sample in a common format and with coordinates corresponding to the same reference genome build.

3.6 Identification of a suitable driver prediction strategy

In each of the publications from which the data were derived, attempts were made to identify genes with potential relevance to disease initiation or progression. This problem was approached using different strategies: identification of significantly mutated genes with [65] and without [68] correction for sample or gene-specific mutation rates; integrative multi-omics analysis[209]; and functional effect prediction[222]. As discussed in the Introduction different methods have their strengths and weaknesses and there is no gold standard: disease relevance suggested by any method should ideally be a starting point for hypothesis generation and experimental investigation. In the absence of a gold standard tool a literature survey of published methods was carried out. Three complementary methods were chosen which were best suited to the characteristics of the dataset, the biological question being explored, and the experimental approach being employed.

To identify significantly mutated genes the widely used MutSigCV[300] tool was chosen. This method was developed specifically to reduce the false-positive rate when deriving SMGs from large cancer sequencing cohorts, so was appropriate for the sample size of 277 used here. It corrects for sample-specific mutation rates, as well as mutation rates varying across different regions of the genome. Because it assesses mutational frequency at the sample level as opposed to the gene level (i.e. x number of samples carry a mutation in gene y, as opposed to gene x carries y number of mutations), its power to detect SMGs scales strongly with cohort size.

To compliment this approach, two gene level prediction algorithms were also selected: OncodriveFM[275] and OncodriveCLUST[276]. These tools were created to address problems with frequency-based approaches so are therefore less depended on sample size. They analyse any gene with at least 2 or 5 mutations, respectively, regardless of how many samples those mutations are distributed across.

OncodriveFM was developed to identify lowly recurrent drivers so is particularly suited for the question at hand. It operates on the assumption that if a gene shows a bias towards harbouring functionally detrimental mutations this indicates that loss of function is being selected for during tumour evolution a pattern indicative of a tumour suppressor gene.

Lastly, OncodriveCLUST was developed to overcome limitations of both frequency-based and functionality-based methods; it focuses instead on spatial clustering of non-synonymous mutations around specific residues along the amino acid sequence. It is therefore best suited for predicting putative neomorphic mutations which often show this pattern of distribution.

Whilst the power and accuracy of the latter methods inevitably increases with cohort size, the metric being assessed for each is not frequency of occurrence as is seen in MutSigCV. For OncodriveFM the relevant derived metric is the ratio of functionally detrimental mutations to functionally benign mutations in a given gene; for OncodriveCLUST the spatial distribution of non-synonymous mutations is compared against a null model of the distribution of synonymous mutations

derived from each gene in the input. With this in mind these three methods were chosen for driver prediction as their strengths complement each other.

The three algorithms can be implemented individually or together within the IntOgen-mutations[298] pipeline. This pipeline provides a wrapper for OncodriveFM and OncodriveCLUST and was developed by the same group; MutSigCV can be integrated into the pipeline as it was here. IntOgen-mutations performs a pre-processing step for OncodriveFM whereby a functional impact score for each mutation is predicted using PolyPhen2[270], but is then corrected using the TransFIC (Transformed Functional Impact Scores in Cancer) algorithm[299]. Briefly, PolyPhen2 and similar functional impact tools are not optimised for cancer and instead infer the impact of a nucleotide change in an evolutionary context by determining its propensity to vary across multiple species alignments. TransFIC attempts to recontextualise mutations into the setting of human cells. It does this firstly by calculating a baseline tolerance to change for each gene family using the distribution of human germline SNPs as a proxy for tolerance to change in the human lineage. If a gene family harbours many naturally occurring SNPs then it is scored as being more tolerant to nucleotide change. Secondly, if a gene has a high degree of paralogy in the human lineage then it is assumed that loss of the gene through somatic mutation is less impactful to the cell as its paralogs may compensate through functional redundancy. These transformed scores serve as input into OncodriveFM in the IntOgen-mutations pipeline.

3.7 IntOgen-mutations uncovers novel candidate drivers in iCCA

As mentioned above, some of the samples have large numbers of mutations and initially these samples were included in the driver prediction analysis; however, upon reviewing the results there were a substantial number of uncharacterised genes (Open Reading Frames, KIAA genes[3 3]), and those which have previously been identified as known false positives due to their length or numeracy in the genome (mucins [MUC4, MUC21]; olfactory receptors [OR10Z1, OR9G1]; ZNF family members [ZNF181, ZNF443, ZNF696, ZNF823], PCLO)[300]. As such the identification and removal of hypermutated samples was carried out. This involved the removal of samples which met 2 criteria: firstly, they harbour a total number of mutations out-with the 3rd quartile of their respective cohorts' mutation count distribution; and secondly, they exhibit a predominant mutational signature associated with a mutator phenotype or exposure to a prominent mutagen.

This resulted in the identification of 27 statistical outliers and the subsequent removal of 19 samples with predominant Alexandrov signatures[302] linked with tabaco exposure, mismatch repair deficiency, activation induced cytidine deaminase (AID) overactivity, or defective polymerase epsilon. These 19 samples represented only 6.4% of total samples yet contributed 43.9% of all mutations and their removal reduced the number of 'knowable' false-positives in subsequent analyses.

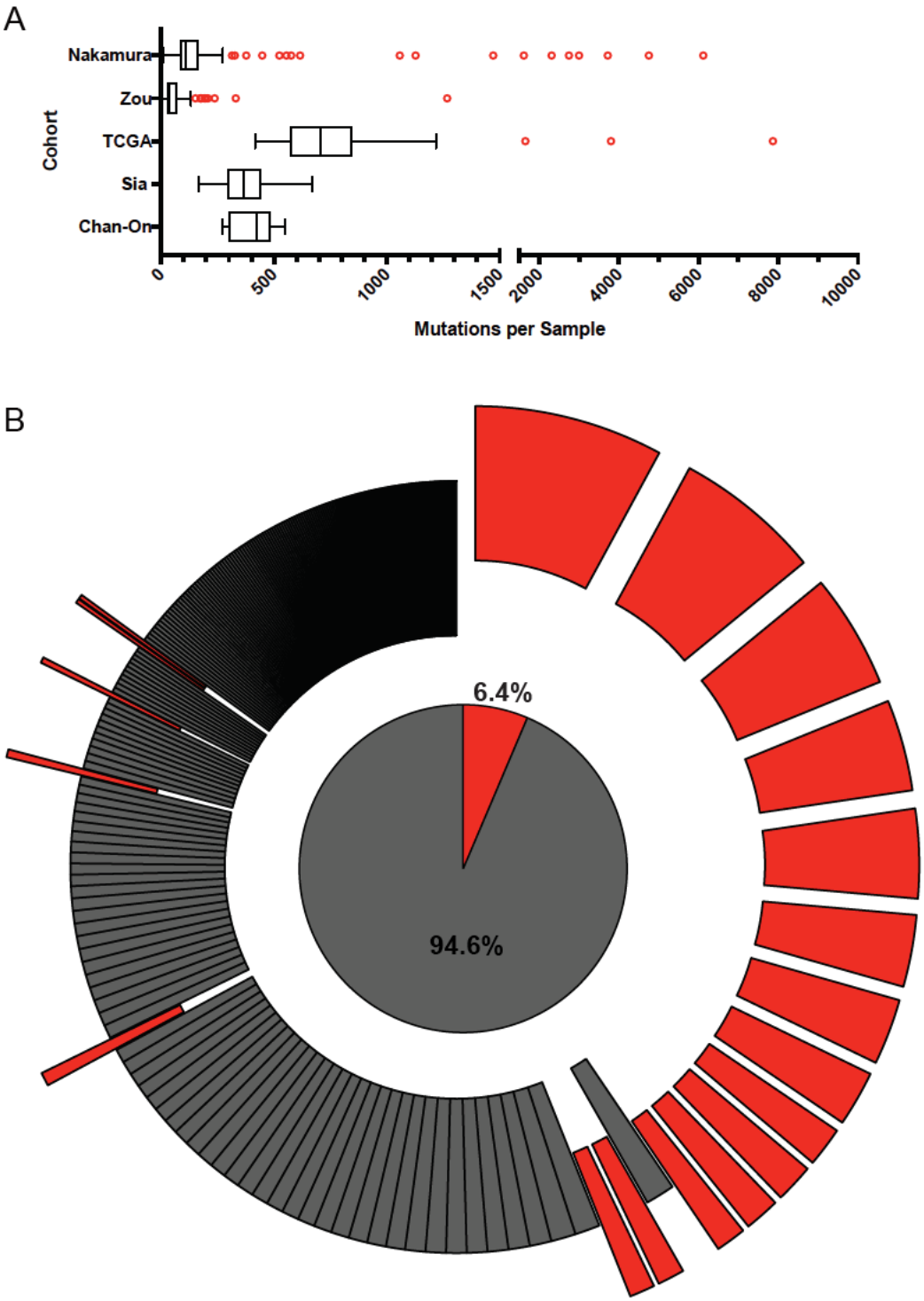

A
Nakamura
Zou
TCGA
Sia
Chan-On
Cohort
0
500
1000
1500
2000
4000
6000
8000
10000
Mutations per Sample
B
6.4%
94.6%

Figure 3.7.1 Identification and Removal of Outliers. A, identification of statistical outliers (red) on a source-by-source basis. The range of mutation count per sample varied between studies and may be attributable to depth of coverage or mutation filtering methods. B, prior to removal 6.4% of samples (inner, red) identified as both statistical outliers and harbouring genomes with Alexandrov signatures 4, 6, 9, or 10 accounted for 43.9% of all mutations (outer, red; segments represent individual samples, size proportional to % contribution of mutations to whole).

Following the removal of hypermutated samples, IntOgen-mutations was run on each constituent cohort (Sia and Chan-On were combined) as well as the pooled cohort. Genes identified in at least one of the analyses were taken forward as candidates, resulting in 96 candidate drivers for experimental screening.

To substantiate the legitimacy of the prediction process, the genes were searched in the COSMIC Cancer Gene Census[3 4,3 5] a curated database of genes for which there is substantial evidence for a functional role in cancer and 32 (33%) were found to overlap. To determine if a similar number of COSMIC genes is likely to be found in a given set of 96 genes by chance alone, 96 genes were randomly sampled from the total set of input genes and the overlap with COSMIC was determined. 100,000 permutations of this resampling were plotted, and it was found that on average a given random set of 96 genes contains only 6 COSMIC genes and that an upper bound of 14 was rarely observed. In a further attempt to quantify this the simulated data from the permutation analysis was used to model the probability of observing at least 32 mutations by random chance in the data as follows: $N(\mu, \sigma)$ is the normal distribution with mean(μ)= 6.26, and standard deviation(σ)=2.32; using this model $P(\geq 32) = 0.01$. It should be noted that this is a rough estimate as it is modelled on a normal distribution which did not accurately fit the simulated distribution (D'Agostino-Pearson, P<0.0001), but more accurate model fitting was out-with the scope of the current analysis and this served only as an approximation. Nevertheless, the above simulations strongly suggest that the enrichment of COSMIC genes is a genuine result of the driver prediction process

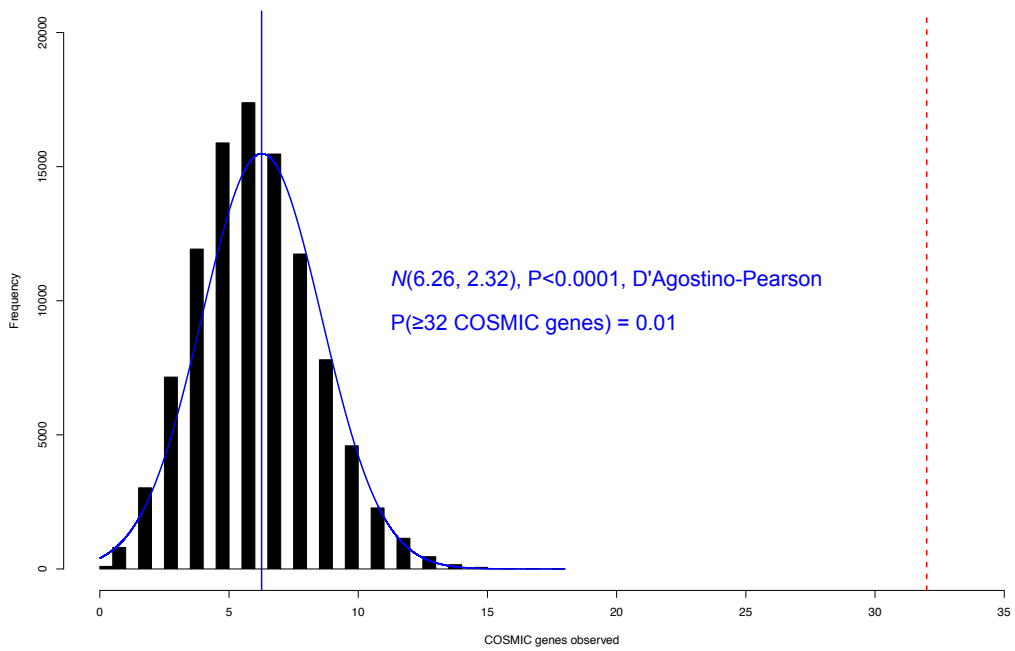

Frequency
0
5000
10000
15000
20000
0
5
10
15
20
25
30
35
COSMIC genes observed
N(6.26, 2.32), P<0.0001, D'Agostino-Pearson
P(≥32 COSMIC genes) = 0.01

Figure 3.7.2 Permutation analysis and statistical modelling of driver prediction. Black bars simulated frequency distribution of 100,000 iterations of the intersect of 96 randomly sampled genes from the input set of genes intersected with the COSMIC database; Blue curve a model of the normal distribution parameterised using the simulated frequency distribution showing the probability of randomly picking 96 genes from the input harbouring at least 32 COSMIC genes (red dashed line) by random chance.

and not due to random chance; this is supported also by the observation that the highest scoring genes are known drivers of iCCA.

Amongst the input of 277 patient samples 55% carried ≥ 2 predicted drivers, and 18% carried none. Regarding the number of samples a predicted driver was found in, it was seen that on average 11 samples harboured a given COSMIC mutation, although this average is skewed by the presence of TP53, KRAS, and ARID1A in this set, which are found at high frequency (n=62, 48, and 33, respectively). Of the predicted drivers which are 'novel' not found in COSMIC they are found on average in 3-4 patients (corresponding to ~1.5% of the sample population) and this result supports the hypothesis that low-frequency drivers may exist in the long tail of infrequently mutated genes found in ICC. Comparing the different driver prediction methods, OncodriveFM predicted significantly more drivers than MutSigCV, however the relative accuracy of each method will be seen during functional validation of this gene set.

Amongst the predicted drivers there are a number of cancer-associated genes that are only predicted by one of the methods. Key examples include NF2, only identified by MutSigCV and only when all samples were combined; known PIK3CA hotspot mutations with clinical relevance (E545 and H1047) [06,3 6] only identified by OncoDriveCLUST; and the tumour suppressor SETD2, a histone lysine methyltransferase identified by OncodriveFM. Interestingly, whilst NF1 mutations have been reported in iCCA[209], the related NF2 have not, despite being shown in models to induce iCCA. Loss of NF2 activates the Hippo pathway, activation of which through other methods have been used to initiate iCCA in murine models of

the disease [76]. Similarly, SETD2 loss has not been explored in iCCA yet given its role in modulating the epigenome[3] [7] it fits in amongst the other epigenetic modulators mutated in iCCA: ARID1A, IDH1/2, BAP1 and PBRM1 indeed the association between mutations in BAP1, PBRM1, and SETD2 have been explored in clear cell renal cell carcinoma[3] [8].

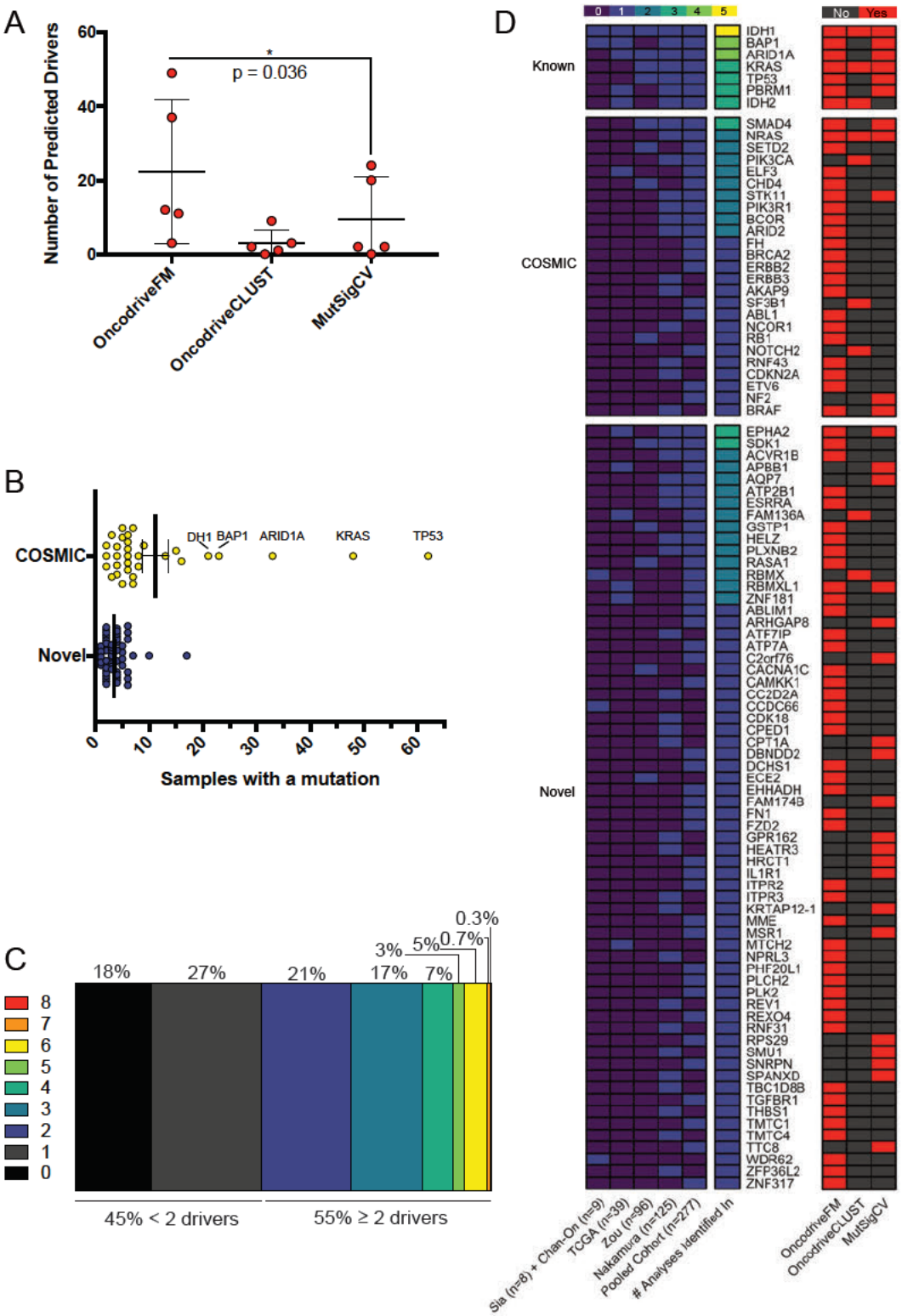

A
Number of Predicted Drivers
60
40
20
0
*
p = 0.036
OncodriveFM
OncodriveCLUST
MutSigCV
B
COSMIC
Novel
DH1 BAP1
ARID1A
KRAS
TP53
0
10
20
30
40
50
60
Samples with a mutation
C
8
7
6
5
4
3
2
1
0
18%
27%
21%
17%
7%
3%
5%
0.7%
0.3%
45% < 2 drivers
55% ≥ 2 drivers
D
0 1 2 3 4 5
No Yes
Known
COSMIC
Novel
IDH1
BAP1
ARID1A
KRAS
TP53
PBRM1
IDH2
SMAD4
NRAS
SETD2
PIK3CA
ELF3
CHD4
STK11
PIK3R1
BCOR
ARID2
FH
BRCA2
ERBB2
ERBB3
AKAP9
SF3B1
ABL1
NCOR1
RB1
NOTCH2
RNF43
CDKN2A
ETV6
NF2
BRAF
EPHA2
SDK1
ACVR1B
APBB1
AQP7
ATP2B1
ESRRA
FAM136A
GSTP1
HELZ
PLXNB2
RASA1
RBMX
RBMXL1
ZNF181
ABLIM1
ARHGAP8
ATF7IP
ATP7A
C2orf76
CACNA1C
CAMKK1
CC2D2A
CCDC66
CDK18
CPED1
CPT1A
DBNDD2
DCHS1
ECE2
EHHADH
FAM174B
FN1
FZD2
GPR162
HEATR3
HRCT1
IL1R1
ITPR2
ITPR3
KRTAP12-1
MME
MSR1
MTCH2
NPRL3
PHF20L1
PLCH2
PLK2
REV1
REXO4
RNF31
RPS29
SMU1
SNRPN
SPANXD
TBC1D8B
TGFBR1
THBS1
TMTC1
TMTC4
TTC8
WDR62
ZFP36L2
ZNF317
Sia (n=8) + Chan-On (n=9)
TCGA (n=39)
Zou (n=96)
Nakamura (n=125)
Pooled Cohort (n=277)
Analyses Identified In
OncodriveFM
OncodriveCLUST
MutSigCV

Figure 3.7.3 IntOgen predicts known and novel drivers. A, OncodriveFM predicts significantly more drivers than MutSigCV. B, the number of samples each driver is found in segregated into COSMIC and novel genes; the average is higher for COSMIC genes, with established iCCA drivers found at high frequency in the cohort. In line with the hypothesis, on average a given novel driver is found in only 3-4 samples. C, the percentage of samples with a given number of drivers, the majority harbour 2-3 drivers, with only a tiny fraction harbouring 7-8. D, the complete set of predicted drivers, whether they are known, found in COSMIC or novel, the analyses they are found in and the prediction algorithm that identifies them in the pooled cohort analysis. A complete set of the mutations in genes identified as drivers can be found in Appendix 4.

3.8 Pathway and Functional Interaction Analysis

To explore the functional significance of predicted drivers, networks were constructed based on known and predicted physical interactions[28,303] and clustered into modules based on connection density[3,9]. This produced a network with 6 modules containing a mix of known, COSMIC, and novel genes. The use of filler genes (those connecting predicted driver genes but not found in the prediction analysis) were not used to construct this network, as direct interactions between components was preferable for biological relevance and clarity. Gene ontology analysis was performed on the modules to ascertain the biological processes each module may participate in. Processes featuring novel drivers, or novel drivers in combination with better known cancer genes include negative regulation of cell growth (APBB1, STK11), negative regulation of cell proliferation and migration (NF2), and regulation of angiogenesis (EPHA2), amongst others. Importantly, the connectivity between novel and known genes in shared processes suggests that although mutated at low frequency in the patient population these novel drivers may contribute to cancer phenotypes by disrupting the same processes as more established cancer genes. Finally, to determine whether or not any of the predicted drivers represent potential new drug targets, known cancer drug-gene interactions were projected onto the network. This revealed ABL as a network hub, however the drugs associated with this gene are largely for the targeting of ABL fusions seen in leukaemia and therefore may be of limited utility in iCCA, where ABL was not seen to participate in oncogenic fusions.

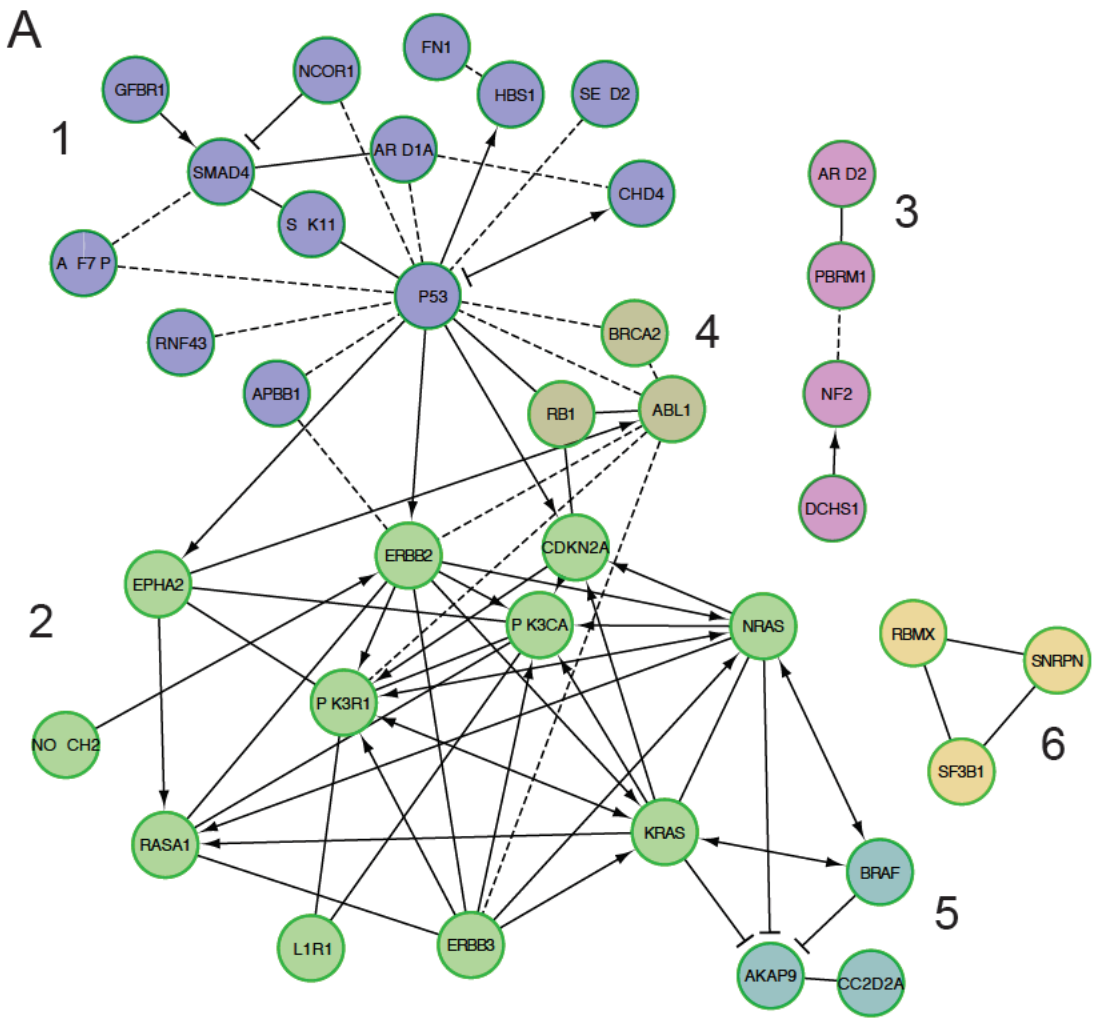

A
1
2
3
4
5
6
FN1
NCOR1
GFBR1
HBS1
SE D2
AR D1A
SMAD4
CHD4
S K11
A F7 P
P53
RNF43
BRCA2
APBB1
RB1
ABL1
AR D2
PBRM1
NF2
DCHS1
ERBB2
CDKN2A
EPHA2
P K3CA
NRAS
RBMX
SNRPN
P K3R1
NO CH2
SF3B1
RASA1
KRAS
BRAF
L1R1
ERBB3
AKAP9
CC2D2A

Figure 3.7.4 Functional interaction network. A, Network constructed from predicted drivers for which evidence of a functional interaction (solid lines, arrow - positive interaction, blunt - negative interaction) is known, or predicted (dashed). The network is clustered into a community of subnetworks termed modules based on density of connectedness. Numbers link a given module to its biological process enrichment in table 3.7.3

Module	Biological Process	P-value	FDR	Genes
1	cell cycle arrest	1.30E-07	4.88E-05	THBS1,TGFBR1,STK11,APBB1,TP53
1	positive regulation of TGFb receptor signaling pathway	2.68E-06	4.83E-04	SMAD4,THBS1,STK11
1	negative regulation of cell growth	3.86E-06	4.83E-04	SMAD4,STK11,APBB1,TP53
1	positive regulation of transcription, DNA-templated	6.81E-06	6.40E-04	ATF7IP,SMAD4,ARID1A,TGFBR1,APBB1,TP53
1	endothelial cell activation	2.85E-05	2.14E-03	SMAD4,TGFBR1
1	negative regulation of transcription by RNA polymerase II	5.53E-05	3.39E-03	ATF7IP,SMAD4,ARID1A,NCOR1,APBB1,TP53
1	positive regulation of SMAD protein signal transduction	6.40E-05	3.39E-03	SMAD4,TGFBR1
2	ERBB2 signaling pathway	8.69E-09	1.12E-06	ERBB3,PIK3CA,ERBB2,PIK3R1
2	phosphatidylinositol 3-kinase signaling	9.90E-09	1.12E-06	ERBB3,PIK3CA,ERBB2,PIK3R1
2	negative regulation of cell-matrix adhesion	3.50E-07	2.62E-05	CDKN2A,RASA1,PIK3R1
2	MAPK cascade	5.46E-07	3.06E-05	NRAS,ERBB3,RASA1,ERBB2,KRAS
2	phosphatidylinositol phosphorylation	1.46E-06	6.59E-05	ERBB3,PIK3CA,ERBB2,PIK3R1
2	negative regulation of neuron apoptotic process	2.85E-06	1.05E-04	ERBB3,PIK3CA,RASA1,KRAS
2	positive regulation of protein kinase B signaling	5.43E-06	1.74E-04	ERBB3,PIK3CA,ERBB2,PIK3R1
2	Ras protein signal transduction	2.00E-05	5.60E-04	NRAS,CDKN2A,KRAS
2	positive regulation of cellular senescence	3.80E-05	8.38E-04	CDKN2A,KRAS
2	positive regulation of phosphatidylinositol 3-kinase signaling	3.81E-05	8.38E-04	ERBB3,PIK3CA,PIK3R1
2	cytokine-mediated signaling pathway	5.76E-05	1.15E-03	PIK3CA,IL1R1,KRAS,PIK3R1
2	bone remodeling	7.05E-05	1.27E-03	NOTCH2,EPHA2
2	transmembrane receptor protein tyrosine kinase signaling pathway	7.91E-05	1.35E-03	ERBB3,ERBB2,EPHA2
2	peripheral nervous system development	1.51E-04	2.42E-03	ERBB3,ERBB2
2	regulation of ERK1 and ERK2 cascade	2.11E-04	3.16E-03	ERBB2,EPHA2
2	regulation of cell motility	2.44E-04	3.42E-03	ERBB3,ERBB2
2	regulation of angiogenesis	2.80E-04	3.46E-03	ERBB2,EPHA2
2	peptidyl-tyrosine phosphorylation	2.88E-04	3.46E-03	ERBB3,ERBB2,EPHA2
2	protein kinase B signaling	4.25E-04	4.67E-03	PIK3CA,EPHA2
3	nucleosome disassembly	1.24E-05	6.57E-04	PBRM1,ARID2
3	negative regulation of cell proliferation	9.26E-05	2.41E-03	PBRM1,NF2,ARID2
3	negative regulation of cell migration	3.48E-04	4.52E-03	NF2,ARID2
4	positive regulation of mitotic cell cycle	1.45E-05	2.50E-03	ABL1,BRCA2
4	cell cycle arrest	3.07E-04	4.54E-03	RB1,ABL1
5	positive regulation of peptidyl-serine phosphorylation	1.21E-04	5.00E-03	AKAP9,BRAF
5	ciliary basal body-plasma membrane docking	1.55E-04	5.00E-03	AKAP9,CC2D2A
6	mRNA splicing, via spliceosome	5.29E-06	8.99E-05	SNRPN,SF3B1,RBMX

Table 3.7.3 Biological process enrichment on a module-by-module bases.

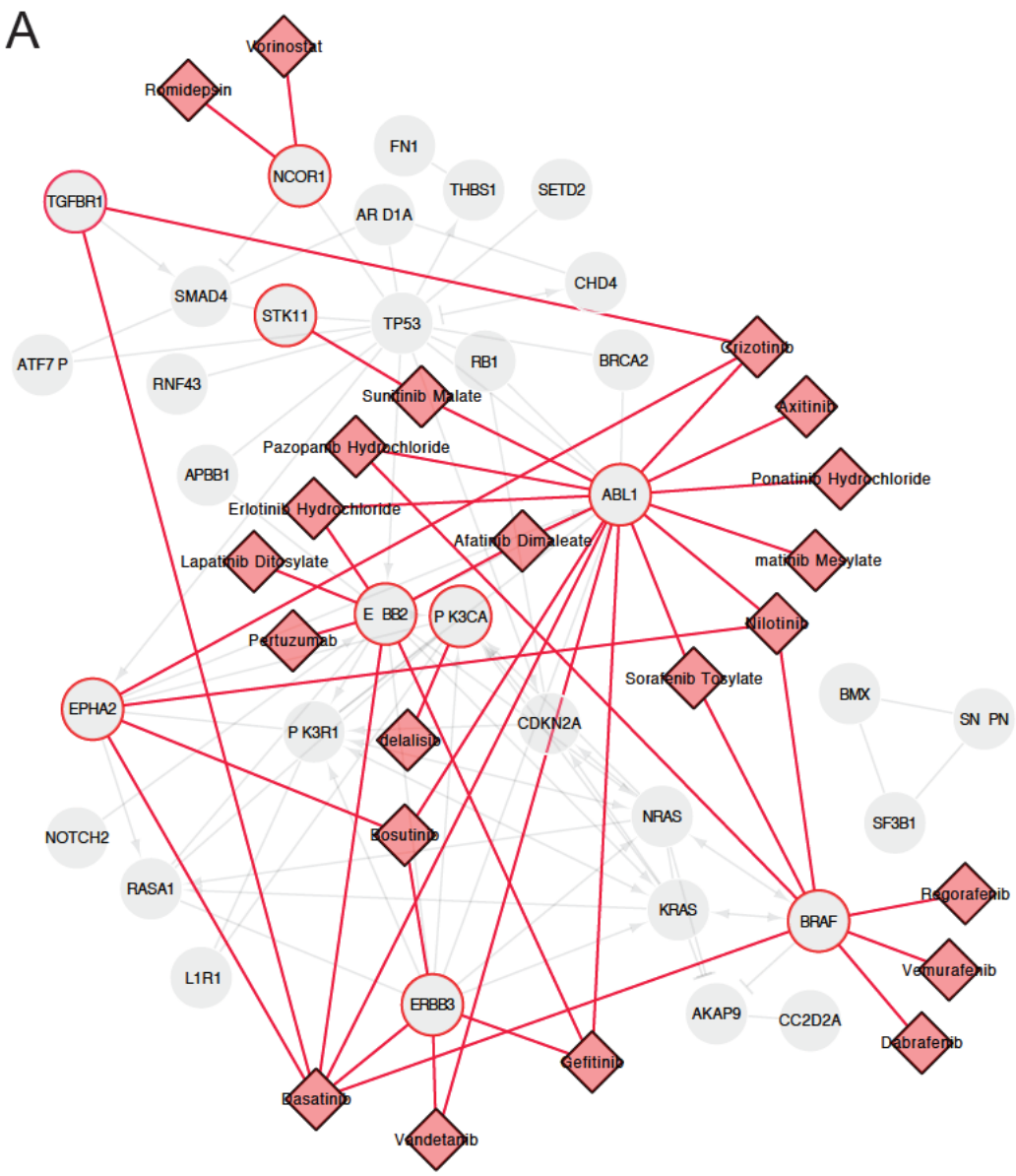

A
Vorinostat
Romidepsin
NCOR1
FN1
THBS1
SETD2
TGFBR1
AR D1A
CHD4
SMAD4
STK11
TP53
Crizotinib
ATF7 P
RNF43
RB1
BRCA2
Sunitinib Malate
Axitinib
Pazopanib Hydrochloride
APBB1
ABL1
Ponatinib Hydrochloride
Erlotinib Hydrochloride
Afatinib Dimaleate
Lapatinib Ditosylate
matinib Mesylate
E BB2
P K3CA
Nilotinib
Pertuzumab
Sorafenib Tosylate
BMX
EPHA2
SN PN
P K3R1
idelalisib
CDKN2A
NRAS
SF3B1
NOTCH2
Bosutinib
RASA1
KRAS
BRAF
Regorafenib
Vemurafenib
L1R1
ERBB3
AKAP9
CC2D2A
Dabrafenib
Gefitinib
Dasatinib
Vandetanib

Figure 3.7.5 Drug target identification. A, The interaction network from 3.7.4 with cancer drugs (red diamonds) projected onto it and their gene targets highlighted with red borders.

3.9 Discussion

The above analysis supports the hypothesis that the long tail of infrequently mutated genes in iCCA contains putative driver genes, and that such genes with disease relevance can be mined from existing data using a rational selection of prediction tools. Existing publications provides a rich source of under-utilised data which can be pooled together to provide the statistical power to detect rare patterns of occurrence that are not visible when considering the component studies in isolation. The union of different prediction tools with complementary strengths and weaknesses, acting on a combined sample set of a size not yet seen in iCCA, revealed a promising set of predicted drivers for experimental screening. Furthermore, genes previously unexplored in iCCA exist in interaction networks alongside known iCCA drivers and established cancer genes generally, and such network participate in disease-relevant processes. The identification of high-confidence novel drivers through the use of functional genomic screening will begin to establish the prediction accuracy of this bioinformatics approach to driver gene discovery.

Chapter 4 CRISPR/Cas9 Editing in the Adult Murine Liver

4.1 Introduction

The Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR)/CRISPR-Associated 9 (Cas9) system has been adapted from the bacterial immune system for molecular biology applications, namely precision genome editing[320,32] . Cas9 encodes an endonuclease capable of cleaving dsDNA to produce a blunt-ended double strand break (DSB). In its endogenous bacterial setting Cas9 is directed to its target cleavage site by dual RNAs constituting a homing CRISPR RNA (crRNA), and a *trans*-activating crRNA (tracrRNA) that is required for maturation of the holoenzyme[322]. In this setting, the crRNA directs the Cas9 enzyme to a target site where Cas9 mediates localised melting of the dsDNA and the crRNA subsequently hybridises with its complementary genomic sequence. Cas9 then cleaves within this genomic target to generate a DSB. Target suitability is constrained by the requirement for a 3-nucleotide conserved protospacer adjacent motif (PAM) next to the target site. In this way Cas9 is programmed by the crRNA and can be directed to almost any region of the genome provided the presence of a proximal PAM site[323].

This programmability has made CRISPR/Cas9 highly effective in experimental systems for precision editing of loci in a variety of genomes. For ease in experimental applications the dual RNA guidance system found in nature has been simplified into a single RNA called a short guide RNA (sgRNA)[322]. Repair of the DSB

follows either the non-homologous end joining (NHEJ) pathway which repairs the blunt-ended DSB with the frequent introduction of frameshifting insertion or deletion (indel) mutations. Alternatively, if an exogenously supplied homologous repair template is provided the DSB can be repaired albeit at low frequency by homology directed repair (HDR). These two repair pathways allow for the targeted knockout of gene expression through frameshifting mutation, or the modification of a locus with carefully designed single or dual sgRNAs and repair templates.

This system is tractable and amenable to multiplexing[286]. These properties make it suitable for high-throughput targeted mutagenesis screening and therefore appropriate for the screening of candidate driver genes. Targeted gene disruption through DSB-NHEJ can identify genes which normally function as tumour suppressors. Alternatively, the inclusion of a template carrying a specific point mutation allows investigation of putative oncogenes through DSB-HDR; however, this is not easily multiplexed and present approaches suffer from low efficiency[324], therefore screening of putative oncogenes was not pursued here and the decision was made to focus on identifying novel tumour suppressors.

A primary obstacle is delivery of CRISPR/Cas9 *in vivo* to a specific target tissue. The delivery of genetic constructs to tissues of interest has been primarily achieved through packaging of DNA constructs in viral (retro-/lenti-, adeno-(associated))[325] or non-viral vectors (lipoplexes/polyplexes)[326] with modification to the packaging to achieve tissue-specific tropism.

For targeting the liver in mice, it is possible to deliver naked DNA in solution due to the anatomy of the liver vasculature[327]. The capillaries of the liver, known as liver sinusoids, have a discontinuous fenestrated endothelium lacking tight junctions and basement membrane which allows facile materials exchange between the vasculature and the parenchyma. The rapid and high-pressure (hydrodynamic) injection of a large volume of DNA solution into the circulation via the tail vein results in cardiac congestion and causes retrograde flow back out of the atrium, through the vena cava and into the liver through the hepatic vein. In the liver, the hydrodynamic force causes enlargement of fenestrae and passage of plasmid solution into the space of Disse surrounding the hepatocytes[328]. The pressure causes deformation of the hepatocyte membrane and the formation of pores through which the plasmid solution enters the cytoplasm - a process termed hydroporation[329]. The pores subsequently close within 10 minutes post-injection and heart rate returns to normal ~1 minute post-injection; mice that have undergone hydrodynamic tail vein injection (HDTVI) experience a life span comparable to uninjected mice.

In this chapter the efficiency of editing *in vivo* by a CRISPR/Cas9 plasmid in the liver following hydrodynamic tail vein injection was investigated. Through the use of an *mT/mG* reporter mouse it was found that this was an efficient method for delivery and target disruption in the adult murine liver, achieving sufficient editing efficiency for candidate driver screening.

4.2 Hypothesis

Targeted genome editing in the adult murine liver is achievable by CRISPR/Cas9 plasmid delivery via hydrodynamic tail vein injection; furthermore, multiple plasmids can be delivered to a single cell.

4.3 Aims

- Validate the *mT/mG* reporter system *in vitro* as a readout of CRISPR/Cas9-induced DSB-NHEJ at a target locus.
- Confirm CRISPR plasmid delivery and editing of the *mT/mG* locus and the endogenous *Apc* locus in the adult murine liver.
- Determine the feasibility of delivering multiple plasmids simultaneously to individual cells *in vivo*.

4.4 The *mT/mG* reporter can be used as a visual readout of CRISPR/Cas9 editing

The duel colour genetic *mT/mG* reporter allele encodes a floxed membrane-tagged tandem dimer Tomato (tdTomato: *'mT'*) and polyadenylation signal, followed by a membrane-tagged Enhanced Green Fluorescent Protein (EGFP: *'mG')* and polyadenylation signal (fig 4.4.1; A). Expression at the locus is driven by a cytomegalovirus enhancer-chicken beta-actin promoter (pCA) and results in expression of the tdTomato fluorescent protein[330]. Upon *Cre*-mediated recombination, the tdTomato cassette including its stop codon are excised and the EGFP fluorescent protein is instead expressed from the pCA promoter following DSB repair. This system has classically been used to investigate questions regarding tissue-specific expression by, for example, expressing *Cre* from a suspected tissue-specific promoter in $R26^{mT/mG}$ mice and assessing the tissue for a red-green colour change[33 ,332]. Here instead the system was used to report the efficiency of CRISPR/Cas9 editing *in vivo* by targeting sgRNAs to loxP sites in the absence of *Cre*[333]. The loxP-targeting sgRNA (sgloxP) guides Cas9 to the loxP sites resulting in the introduction of DSBs and excision of the tdTomato and stop codon. Subsequent repair by NHEJ results either in EGFP expression, or loss of fluorescence if the repair shifts the EGFP expression cassette out of reading frame (fig 4.4.1; C). If generation and delivery of a plasmid carrying Cas9 and sgRNA expression cassettes was successful, then liver cells which have received and expressed the plasmid show a loss of tdTomato expression with or without a colour change to green.

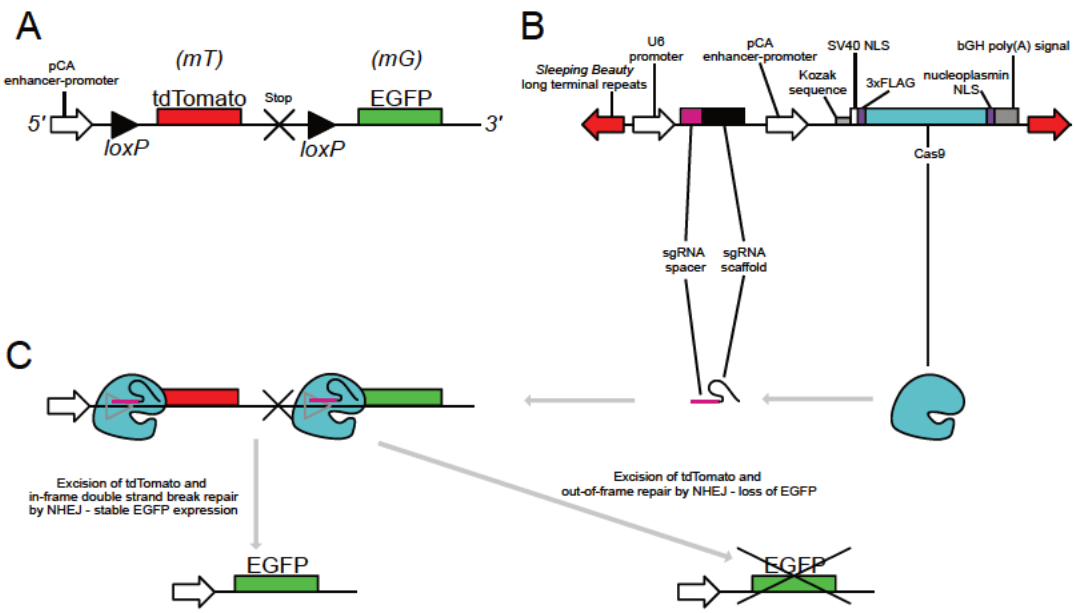

A
pCA
enhancer-promoter
(mT)
tdTomato
Stop
(mG)
EGFP
5′
loxP
loxP
3′
B
Sleeping Beauty
long terminal repeats
U6
promoter
pCA
enhancer-promoter
Kozak
sequence
SV40 NLS
3xFLAG
nucleoplasmin
NLS
bGH poly(A) signal
Cas9
sgRNA
spacer
sgRNA
scaffold
C
Excision of tdTomato and
in-frame double strand break repair
by NHEJ - stable EGFP expression
EGFP
Excision of tdTomato and
out-of-frame repair by NHEJ - loss of EGFP
EGFP

Figure 4.4.1 mTmG reporter locus and SB-CRISPR plasmid structure. A, schematic of the mT/mG reported allele: the mT element is flanked by loxP sites and followed by the mG element. Whist the mT element is present, EGFP is not expressed from the CMV enhancer-chicken beta-actin promoter (pCA). B, schematic of the SB-CRISPR plasmid: a sgRNA spacer-scaffold sequence is driven by the U6 promoter, the gRNA spacer is flanked by restriction enzyme sites (not shown) to enable facile cloning of 20bp sgRNA spacer sequences targeting desirable genomic loci. FLAG-tagged Cas9 is flanked by nuclear localisation signals (NLS) and driven by the same pCA promoter as is found in the mT/mG reporter. Sleeping Beauty LTRs facilitate integration of the CRISPR construct into the host genome. C, Editing schematic for sgloxP in mT/mG reporter cells. Upon expression of sgloxP and Cas9 from the SB-CRISPR plasmid, sgloxP and Cas9 form the catalytically active Cas9 holoenzyme which is directed to the mT/mG loxP sites. Editing and subsequent double-strand break repair by non-homologous end joining (NHEJ) results in excision of the mT element and either EGFP expression from the pCA promoter, or loss of fluorescence if the repair process shifts EGFP out of reading frame.

To confirm the utility of both the reporter and the sgloxP guide before testing *in vivo*, an *in vitro* editing experiment was first performed in murine bile duct organoids derived from the *R26*$^{mT/mG}$ mouse, homozygous for the reporter allele. Two lines were used: one stably expressing human codon-optimised *Streptococcus pyogenes* Cas9 (referred to as Cas9 from now onwards) and a second stably expressing a catalytically inactive Cas9 ('dead Cas9': dCas9). Either A 20bp spacer targeting the loxP locus (sgloxP), or a random 20bp sequence with no known genomic target (sg0007), was cloned into the sgRNA expression cassette of the LentiGuide-Puro (LGP) plasmid[287]. This was subsequently packaged as a lentivirus with broad tissue tropism through pseudotyping with the vesicular stomatitis virus envelope glycoprotein (VSV-G)[334]. 36 hours following transduction of Cas9 and dCas9 organoid fragments with either LGP-sgloxP or LGP-sg0007, the resulting organoids surviving puromycin selection were imaged either on a confocal microscope (Cas9), or a fluorescence light microscope (dCas9) (fig 4.4.2; A, B, respectively).

In Cas9 organoids that received sgloxP there were several types of colour change observed: where cells were homozygously edited, both red-to-green and red-to-colourless is observed (fig 4.4.2; A, middle panel). When only one copy of the locus was edited there was expression of both EGFP and tdTomato resulting in yellow upon merging channels (fig 4.4.2; A, right panel). There may also be heterozygous editing in cells which appear red due to loss of EGFP reading frame in a single successfully edited *mT/mG* locus, but this outcome cannot be resolved through imaging alone and sequencing of the locus in the different organoid populations

was not carried out. Mosaicism is seen in some organoids which have successfully been edited because lentiviral transduction was carried out on organoid fragments, not single cells (fig 4.4.2; A, middle panel). If not all cells in the fragment were transduced, then the resultant organoids were mosaic; this mosaicism is lost with longer selection time and multiple passages (data not shown).

dCas9-sgloxP (fig 4.4.2, A, left panel), as well as Cas9-sg0007 and dCas9-sg0007 organoids showed no sign of red-to-green colour change, indicative of unsuccessful editing at the *mT/mG* loci (fig 4.4.2; B). These observations are in agreement with what is known about Cas9's mechanism of editing. Cas9 requires a sgRNA for its endonuclease function: in its unbound state it is catalytically inactive, only acquiring its DNA recognition and cleavage activities upon sgRNA binding through conformational change[323]. This suggests that in Cas9-sg0007 organoids the lack of a colour change is not attributable to lack of Cas9 endonuclease activity, but to incorrect targeting by the sg0007 guide. In the case of dCas9-sgloxP the lack of editing is likely due to the catalytic inactivity of dCas9 despite binding of an accurately targeting sgRNA. Cas9 cleavage is performed by 2 distinct nuclease domains: an HNH-like nuclease domain (aa1-56) responsible for cleavage of the strand which hybridises to the sgRNA, and a RuvC-like nuclease domain (aa780-906) which cleaves the opposite strand. The dual mutations D10A and H840A found in dCas9 inactivate both of the nuclease domains rendering the enzyme 'catalytically dead' but still able to interrogate and bind target loci. In the case of dCas9-sg0007 the lack of colour change is a product of both a non-targeting sgRNA and catalytically inactive Cas9.

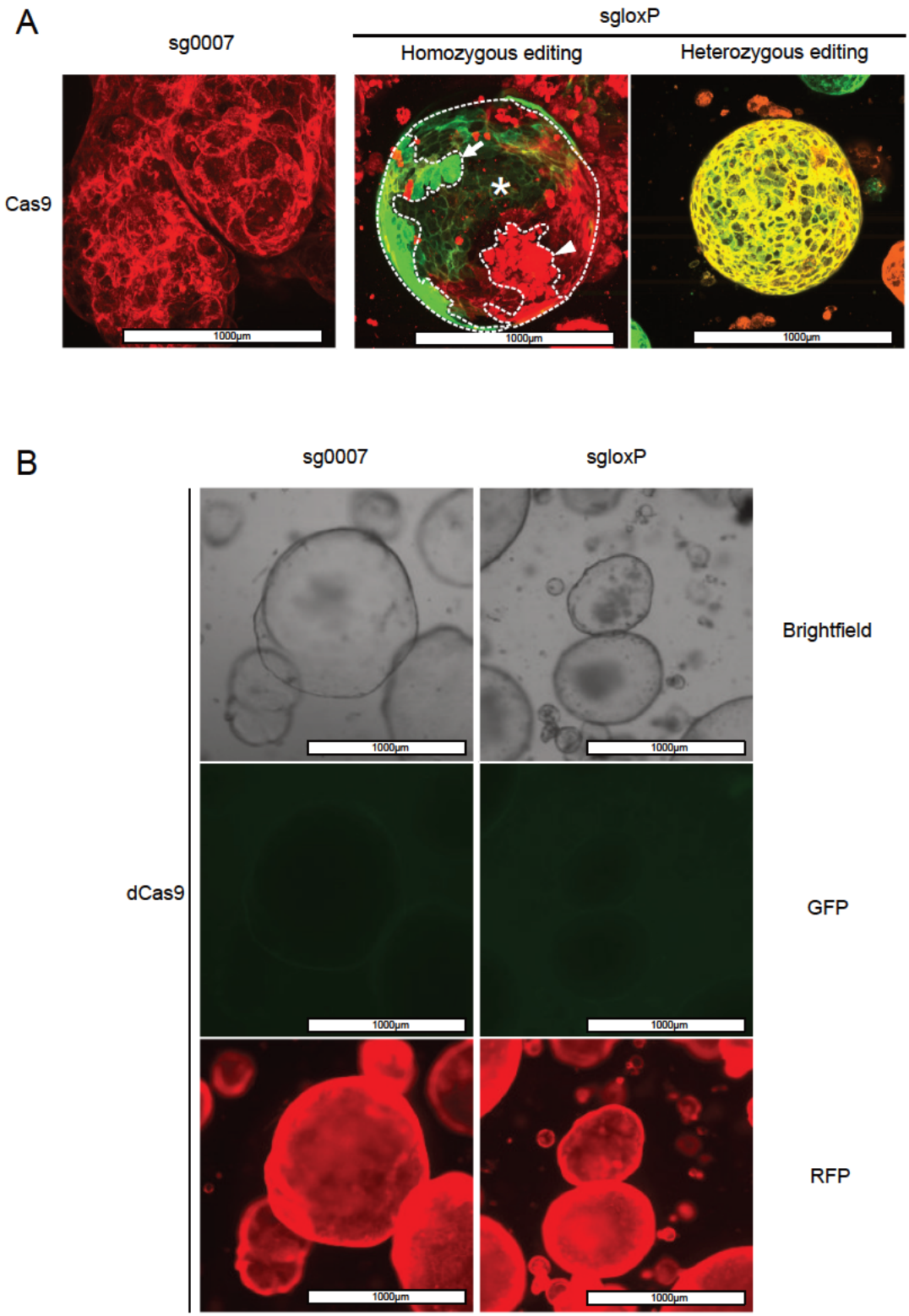

A
sg0007
sgloxP
Homozygous editing
Heterozygous editing
Cas9
1000μm
B
sg0007
sgloxP
dCas9
Brightfield
GFP
RFP
1000μm

Figure 4.4.2 Editing of the mT/mG reporter allele in R26mT/mG biliary organoids. A, in Cas9 expressing organoids transduction of the LentiGuide-Puro plasmid expressing sg0007 results in no colour change, when Cas9 organoids are transduced with sgloxP different colour changes can be seen: homozygous editing (middle panel) results in loss of fluorescence (asterisk) or EGFP expression (arrow), mosaicism is a result of patchy transduction in the organoid fragment producing regions with no editing (arrowhead). Homozygous editing (right panel) is seen when only one locus is edited to produce EGFP expression. B, organoids expressing catalytically inactive dCas9 show no editing at the reporter regardless of sgRNA.

The above results establish that the *mT/mG* reporter system can be used as a visual readout of successful CRISPR/Cas9 targeting, cleavage, and repair by NJEH, and that the components of the system and the protocols to generate them were viable (see Materials and Methods).

4.5 Tissue-specific CRISPR/Cas9 delivery to the adult murine liver: *mT/mG* editing

To determine the efficiency of CRISPR/Cas9 construct delivery to, and editing in, cells in the adult mouse liver via hydrodynamic tail vain injection (HDTVI), the sgRNAs from above were cloned into a *Sleeping Beauty* transposon-CRISPR/Cas9 plasmid (SB-CRISPR)[286] (fig 4.4.1, B). An HDTVI dose titration of SB-CRISPR-sgloxP was carried out in the *R26*$^{mT/mG}$ mouse to determine the optimal dose for delivery and editing. Mice were culled 7 days post-injection and slides were immunohistochemically stained for GFP. This showed that in the presence of transposase (SB13) the number of editing events increased with the concentration of SB-CRISPR-sgloxP over a range of 0.2-20μg when injection volume remained constant. Further, when SB13 was not included with the highest sgloxP concentration, the number of editing events was significantly decreased (fig 4.5.1; A, B). This suggests that with co-delivery of SB13, the SB-CRISPR transposon is integrated and expressed stably as opposed to the CRISPR/Cas9 expression cassettes being expressed transiently from the plasmid without long-term integration, as previously suggested[286].

Focusing on the 20μg + SB13 group there is a range of 0.14-0.67% edited cells, which represents 630-3089 cells in a section of the left lobe of the mouse liver. To arrive at a rough approximation of the number of edited cells in a whole liver; consider that Sohlenius-Sternbeck[335] calculated a mean of 135 ± 10 million (mean ± SEM) hepatocytes per gram of murine liver, and Nascimento-Sales and colleagues[336] notes that the mean liver weight of the FVB strain (the genetic background of the

R26$^{mT/mG}$ used here) on a conventional chow diet is 1.4g ± 0.04 (mean ± SEM). This gives us a rough hepatocellularity number of 189 million cells per FVB liver. If a given section shows 0.14-0.67% editing, then it can be estimated that between 264,600 and 1,266,300 cells (0.14-0.67% of 189 million) may be edited across the whole liver in a given HDTVI with 20μg SB-CRISPR co-delivered with SB13 transposase.

The main factor influencing hepatocellular uptake of the plasmid, other than plasmid concentration, is the hydrodynamic force applied during injection. Optimal delivery is achieved when a solution 10% of the mouse's circulating blood volume is delivered over a 7-second period. This is difficult to keep constant from mouse-to-mouse; total volume delivered, the period over which it is delivered, and the exact proportion of the mouse's circulating blood volume that the injection volume represents will vary slightly between replicates. This may explain the numerical range of editing events seen in each group. However, even at the lower end observed in the 20μg+SB13 group, enough cells are edited to correctly power even a whole-genome screen. For example, there are 129,211 unique sgRNA sequences in the commonly used whole-genome GeCKOv2 library[288]. If we assume that at least one sgRNA sequence is present in every transfected cell, then every unique sgRNA would be represented in at least 2 cells. At the upper range observed in this experiment this number increases to 9 cells per sgRNA.

Although these are approximations which rely on certain assumptions e.g. each lobe and liver region experiencing the same delivery and uptake dynamics this experiment provides evidence that a 20μg HDTVI with SB-CRIPSR in the presence

of SB13 transposase is sufficient to power a sub-genome-scale screening experiment in the adult murine liver.

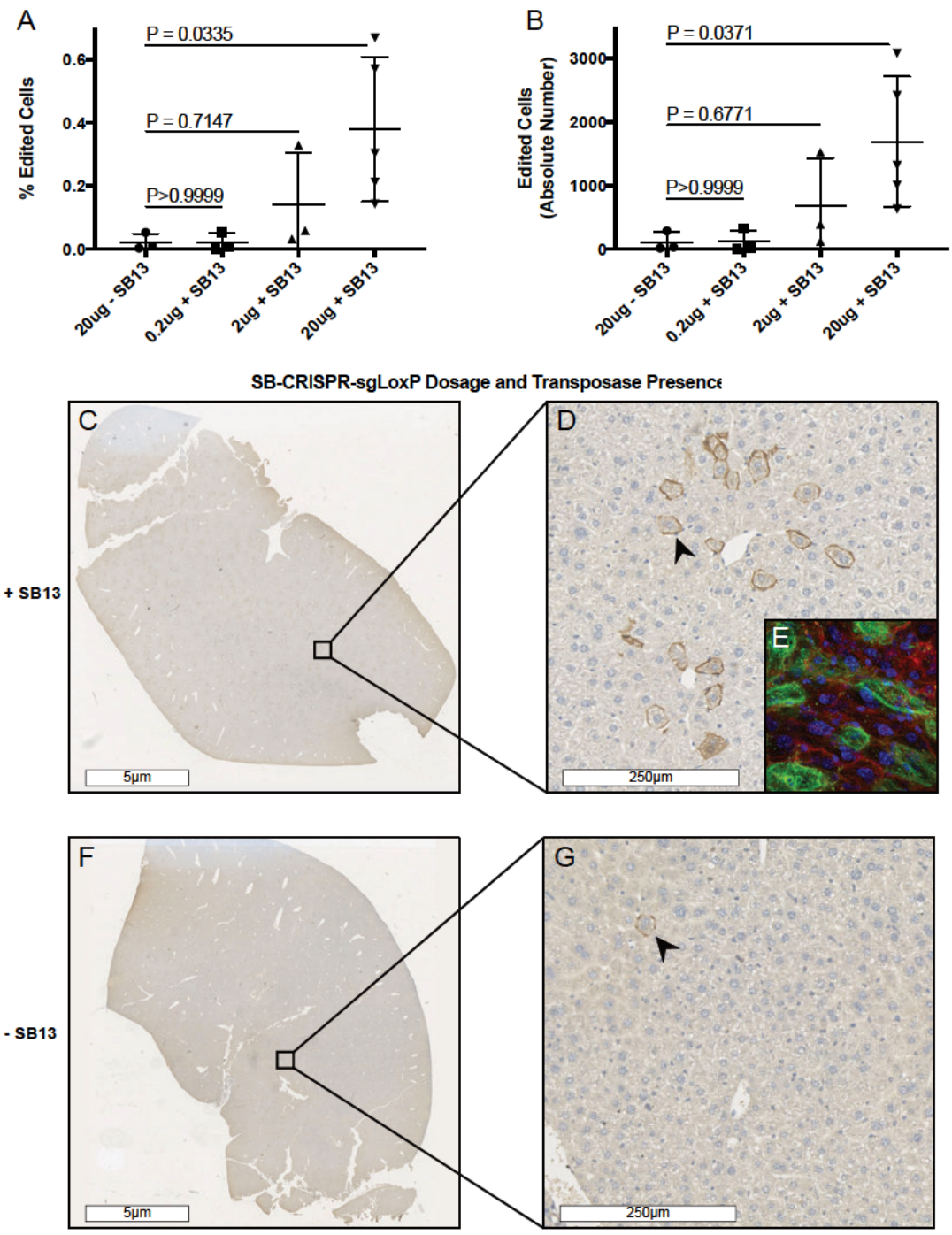

A
% Edited Cells
0.6
0.4
0.2
0.0
P = 0.0335
P = 0.7147
P>0.9999
20ug - SB13
0.2ug + SB13
2ug + SB13
20ug + SB13
B
Edited Cells (Absolute Number)
3000
2000
1000
0
P = 0.0371
P = 0.6771
P>0.9999
20ug - SB13
0.2ug + SB13
2ug + SB13
20ug + SB13
SB-CRISPR-sgLoxP Dosage and Transposase Presence
+ SB13
C
5μm
D
250μm
E
- SB13
F
5μm
G
250μm

Figure 4.5.1 Dose titration of SB-CRISPR-sgloxP. A, the percentage of edited cells at each concentration of SB-CRISPR-sgloxP with and without the presence of SB13. B, the same data as in A, but expressed as number of edited cells. C-E. immunohistochemistry (IHC) for GFP in a liver section from a mouse that received 20μg + SB13. Editing can be seen in small clusters (D), and GFP expression was confirmed by native fluorescence (E). F-G, IHC for GFP in a liver section from a mouse that received 20μg without SB13: editing still occurs but is rarely observed and is presumably the result of transient expression from the SB-CRISPR-sgloxP plasmid.

4.6 Tissue-specific CRISPR/Cas9 delivery to the adult murine liver: *Apc* editing

The above experiment demonstrates the relationship between SB-CRISPR concentration and number of editing events using a convenient genetic reporter. However, an effective screening experiment requires the targeted disruption of endogenous genes resulting in a cell-physiological change. To assess the functional impact of editing an endogenous gene, the *Apc* gene was knocked down using a sgRNA targeting its 1^{s} exon. *Apc* is a component of the canonical Wnt-β-catenin signalling pathway where it functions as a scaffold for the destruction complex. In the absence of Wnt ligand the constitutively produced Wnt effector protein β-catenin is sequentially phosphorylated by components of the destruction complex and subsequently ubiquitinated and degraded by the proteasome. Upon activation by a Wnt ligand, the destruction complex is sequestered to the cell membrane allowing cytosolic β-catenin to evade destruction and translocate to the nucleus. Here it recruits co-activators and transcription factors and initiates a range of transcriptional programs involved in processes such as proliferation, stemness, and epithelial-mesenchymal transition [28,337]. The targeted knockdown of *Apc* causes loss of the destruction complex and allows constitutive activation of Wnt-β-catenin downstream effector signalling: the phenotypic readout of this is increased nuclear β-catenin and increased expression of target genes. Processing of β-catenin by the destruction complex requires an intact N-terminus; the absence of the N-terminus in the ΔN90-β-catenin mutant results in stabilisation of the protein and

increased translocation to the nucleus[338]. In this way ΔN90-β-catenin phenocopies *Apc* loss and provides a suitable positive control.

20μg of either SB-CRISPR-sg*Apc*, SB-CRISPR-sg0007, or pT3-N90-beta-catenin was delivered via HDTVI alongside SB13, and livers were examined 7 days later. Immunohistochemically, there is no increase in nuclear β-catenin in the SB-CRISPR-sg*Apc* group compared to the negative control, whilst the ΔN90-β-catenin positive control group showed a significant increase in nuclear β-catenin (fig 4.6.1; A, B). The lack of nuclear β-catenin in the sg*Apc* group was unexpected, so Wnt pathway activation was checked at the transcriptional level. Via RT-qPCR, it can be seen that β-catenin transcriptional targets are upregulated relative to the non-targeting control, but not to the extent seen in the ΔN90-β-catenin group (fig 4.6.1; C). These data suggest that *Apc* is being knocked down in the sg*Apc* group, but the loss of *Apc* does not produce as strong a phenotype as stabilisation of β-catenin. A possible biological explanation is that the stabilised β-catenin remains in the nucleus for longer and accumulates, resulting in stronger activation of target genes. Conversely, a methodological explanation is that there is simply more β-catenin in the cell in the ΔN90-β-catenin group; the ΔN90-β-catenin protein is surplus to the endogenous copies of wild-type β-catenin and there may be many copies of the expression vector in each cell. The strong nuclear β-catenin staining seen in the ΔN90-β-catenin group supports the hypotheses that there is more cellular β-catenin overall, and/or it is remaining in the nucleus for longer. However, the RT-qPCR data support the successful targeted loss of *Apc in vivo* through CRISPR/Cas9-mediated editing in the liver.

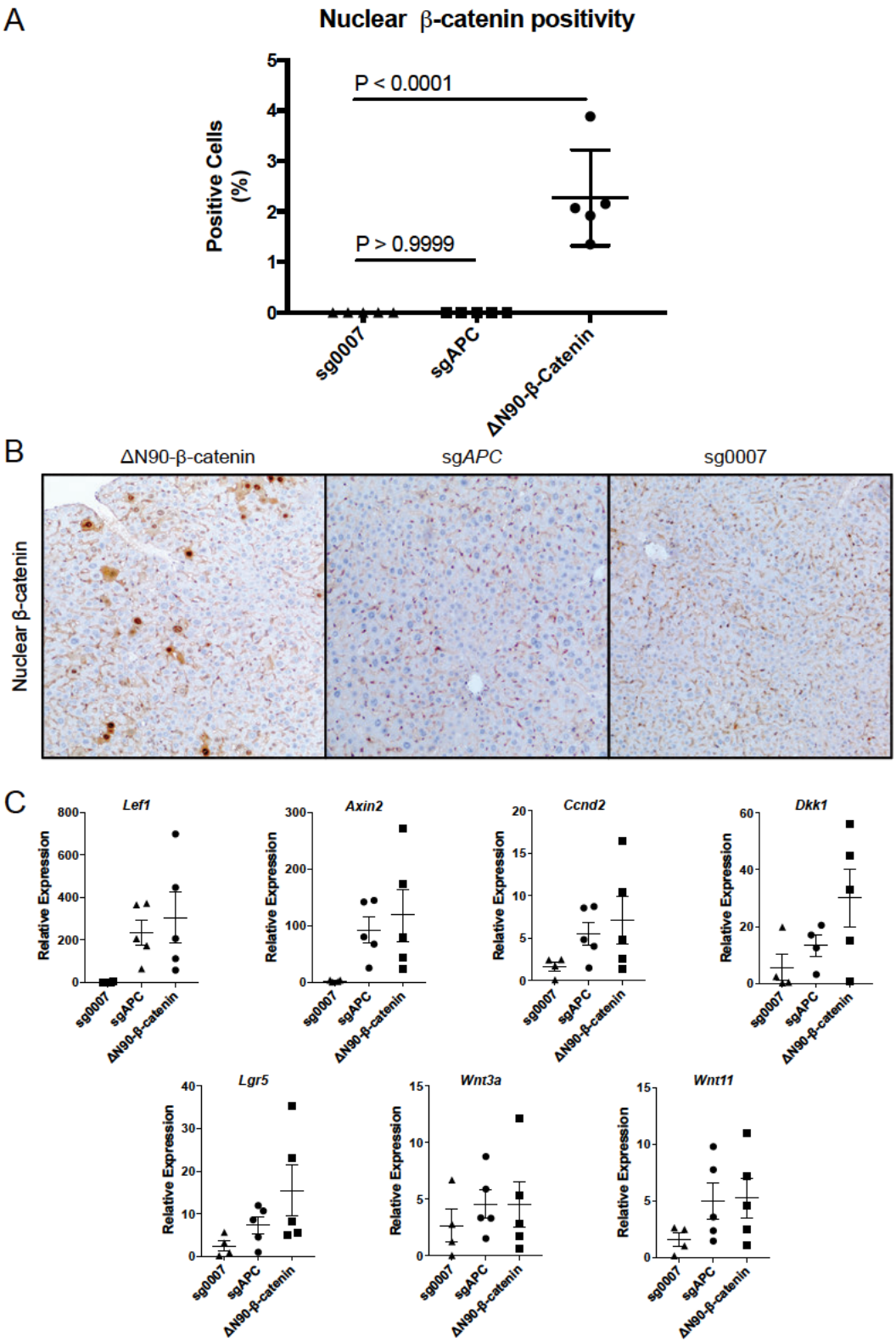

A
Nuclear β-catenin positivity
Positive Cells (%)
P < 0.0001
P > 0.9999
sg0007
sgAPC
ΔN90-β-Catenin
B
ΔN90-β-catenin
sgAPC
sg0007
Nuclear β-catenin
C
Lef1
Axin2
Ccnd2
Dkk1
Lgr5
Wnt3a
Wnt11
Relative Expression
sg0007
sgAPC
ΔN90-β-catenin

Figure 4.6.1 APC editing *in* vivo. A, percentage of cells with nuclear β-catenin positivity by IHC; mice receiving SB-CRISPR-sg0007/sgApc showed no positivity, whilst mice receiving ΔN90-β-catenin had significantly more cells with nuclear positivity ($P<0.0001$, one-way ANOVA). B, representative images of β-catenin IHC quantified in A. C, RT-qPCR results showing expression of β-catenin transcriptional targets relative to GAPDH; mice receiving the SB-CRISPR-sgApc plasmid showed expression intermediate between the negative and positive controls suggestive of endogenous Apc knockdown.

Together with the *mT/mG* editing experiments above, this shows that it is possible to generate and deliver the SB-CRISPR plasmid into an experimentally useful number of cells, that the plasmid integrates and is functional, and this results in the disruption of target genes.

4.7 Delivery of multiple plasmids to individual hepatocytes in a tumour model of intrahepatic cholangiocarcinoma

The induction of tumourigenesis commonly requires more than one genetic alteration[339] [34] . It follows that for effective screening, delivery of multiple SB-CRISPR plasmids is required to induce a number of different editing events; alternatively, delivery of SB-CRISPR screening plasmids alongside oncogene expression vectors is required. In either case multiple distinct plasmids need to be delivered to individual cells. Here an established experimental mouse model of intrahepatic cholangioncarcinoma[99] is used to demonstrate that up to 4 plasmids can be delivered simultaneously to single cells.

The expression of the NOTCH1 receptor intracellular domain (NICD) and constitutively active myristoylated Akt (myrAKT) has been shown to induce iCCA in the murine liver from a hepatocyte origin. The exact mechanism is incompletely understood, however evidence suggests that NICD first induces trans-differentiation of mature hepatocytes into biliary-like cells, and subsequently induces the formation of invasive cystadenocarcinomas when expressed alone[99,342,343]. AKT, when expressed alone gives rise to mainly hepatocellular carcinomas with occasional benign cholangiocellular lesions. When expressed together in mature hepatocytes, lineage switching and subsequent or concurrent transformation is thought to be primarily driven by NICD and accelerated by myrAKT which provides metabolic support to drive proliferation[99].

Using this model alongside the *R26*$^{mT/mG}$/SB-CRISPR-sgloxP system described above, separate plasmids expressing NICD, myrAKT and SB13 were delivered alongside SB-CRISPR-sgloxP to the liver via HDTVI. This resulted in tumour formation after 6 weeks. As observed in the biliary organoid experiments, numerous colour changes were observed reflecting the various zygosity of editing. Predominantly the tumours observed were of a single colour, indicative of tumours of a clonal origin with homozygous editing (fig 4.7.1). There was, however, a single mass with different coloured regions (fig 4.7.2).

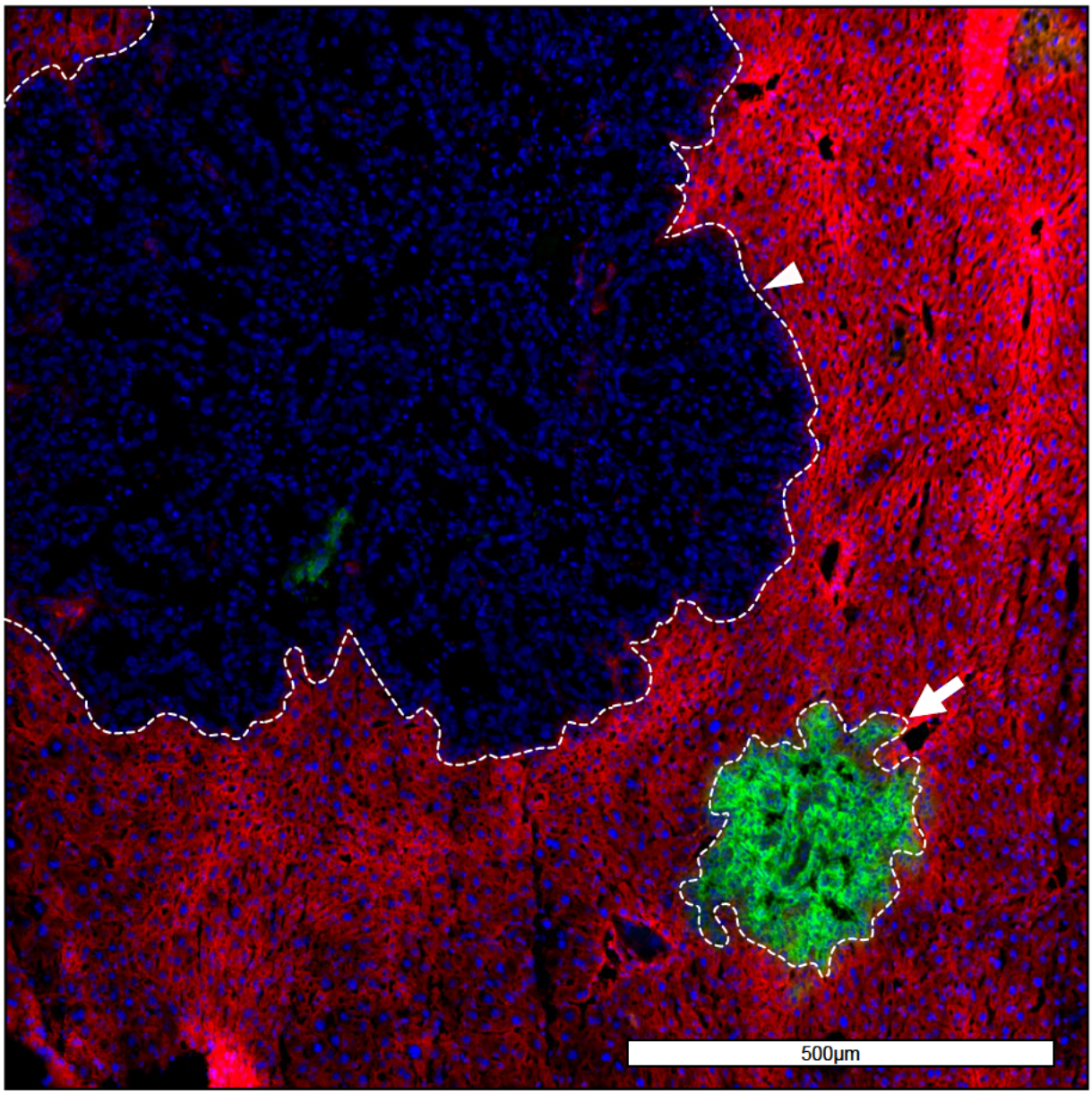

500µm

Figure 4.7.1 clonal editing in mice. $R26^{mT/mG}$ mice injected with plasmids carrying NICD, myrAKT, SB-CRISPR-sgloxP and SB13 expression vectors develop tumours with colour change, suggesting the uptake of all 4 plasmids. Homozygous editing is seen here with two different outcomes as shown in preceding sections. Both loss of fluorescence (arrowhead) and EGFP expression (arrow) is seen in tumours with close proximity. This suggests a clonal origin of these tumours where the respective founder cell has received all plasmids.

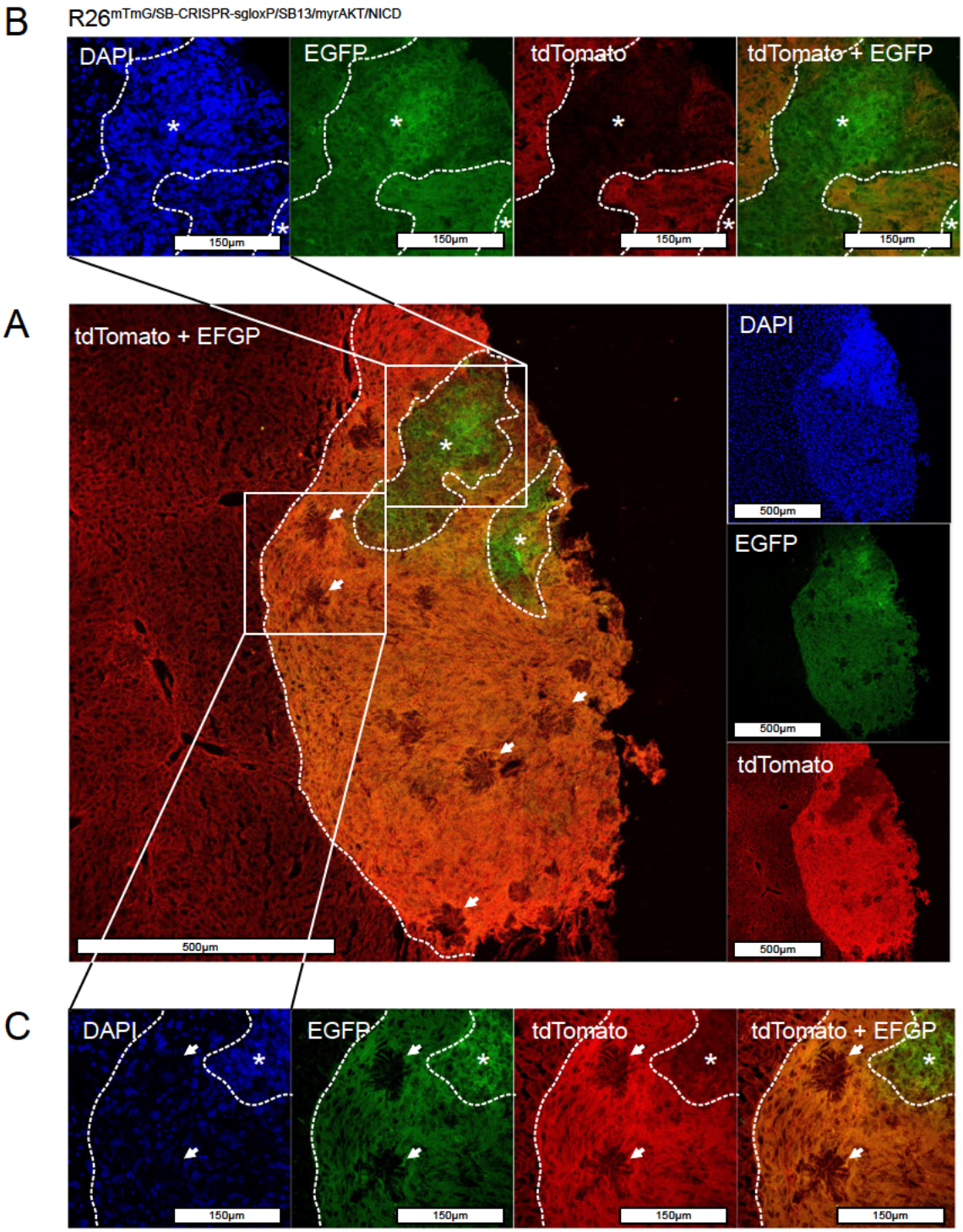

B
R26mTmG/SB-CRISPR-sgloxP/SB13/myrAKT/NICD
DAPI
EGFP
tdTomato
tdTomato + EGFP
150µm
A
tdTomato + EFGP
DAPI
EGFP
tdTomato
500µm
C
DAPI
EGFP
tdTomato
tdTomato + EFGP
150µm

Figure 4.7.2 A, Polyclonal editing in mice. A, an example of a tumour with different mT/mG editing outcomes indicative of a polyclonal and/or polyphyletic tumour. The predominant bulk of the tumour expresses both tdTomato and EGFP, with patches of only EGFP (asterisks) and patches with no fluorescence (arrows). B, detail of an EGFP expressing region which has lost tdTomato expression but is surrounded by tissue expressing both tdTomato and EGFP (merge, right panel). C, detail of two regions with loss of fluorescent, surrounded by tdTomato and EGFP expressing tissue. C, far right panel shows surrounding normal liver with tdTomato expression, and tumour expressing EGFP (asterisk), no fluorescence (arrows), and both tdTomato and EGFP (enclosed in dashed line).

The majority of cells in this tumour express both EFGP and tdTomato as evidenced by comparison with the surrounding tdTomato-expressing unedited liver parenchyma; this is indicative of heterozygous editing of the *mT/mG* reporter. The tumour has regions where only EGFP is expressed, suggesting homozygous editing and in-frame DSB repair; but also has regions which appear to have lost any expression of fluorescent protein suggesting homozygous out-of-frame DSB repair. The different colour changes observed raises the question of whether this tumour arose from a single heterozygously edited cell and underwent clonal diversification with some clones losing the second copy of tdTomato through subsequent editing to produce green clones. This may be possible, however this explanation wouldn't account for the regions without fluorescent protein expression. The locus can only be edited once as excision of the *mT* locus removes the *loxP* target. If the locus were heterozygously edited in the founder cell to produce a clone with both EGFP and tdTomato expression, there could not be a subsequent editing event that results in loss of fluorescence, only an editing event on the unedited locus which would produce a clone with EGFP only. It is therefore more likely that this tumour is polyphyletic arising from multiple individual cells which have different editing outcomes at the *mT/mG* locus. This is further supported by the observation that at this concentration of SB-CRISPR, cells are frequently edited in close proximity (fig 4.5.1; D,E, and fig 4.7.1).

The presence of tumours with colour change suggests that all 4 plasmids were delivered to cells. This supports the proposed mechanism of uptake being non-selective whereby injection of an admixture of plasmids in solution results in

uptake of a random mixture of plasmids in a given cell. When injecting a mixture with a small number of different components it is presumed that in the majority of cells all of the components will be taken up this is observed here as there were no tumours without a colour change, and no normal cells with a colour change. Almost all tumours were either completely EGFP-expressing, or completely colourless demonstrating that in the majority of founder cells homozygous editing was the predominant editing outcome. The tumour in figure 4.7.2 was the only tumour seen with homozygous editing and serves as an important example showing that this mode of editing is a possible if rare outcome, and further shows that tumours arising from HDTVI may represent distinct tumours which have merged. Both of these observations are important when analysing and interpreting any tumours which have arisen from *in vivo* driver screening.

This experiment shows that the delivery and expression of multiple distinct plasmids is possible. What is not known is the relative copy number of each plasmid, or the absolute number of plasmids which are taken up by each cell. It shows that for small plasmid admixtures if a cell receives one plasmid, it likely receives all the different plasmids in the admixture. Although not explored here, there presumably exists a threshold of plasmid uptake above which the uptake is toxic. However, knowledge of the upper limit of plasmids which can be received by a cell is not as important here as knowing that a small number can be delivered. It is thought that between 2-8 driver genes are needed to initiate tumorigenesis depending on the cancer type, and various models of ICC demonstrate that 2 mutations are frequently sufficient. The above experiment shows that at least 4

distinct plasmids can be delivered, therefore a sufficient number of different SB-CRISPR plasmids carrying sgRNAs targeting different candidate genes could be delivered to induce tumourigenesis.

4.8 Discussion

The targeted knockdown of candidate driver genes *in vivo* necessitates tissue-specific delivery of CRISPR/Cas9 at high enough efficiencies to enable representation of the screening library. The above experiments demonstrate that targeted mutagenesis at reporter and endogenous loci is possible through hydrodynamic tail vein injection of CRISPR/Cas9 plasmids, and that delivery of multiple plasmids is achievable in this context. Editing efficiencies were high enough to enable theoretical representation of a whole-genome screening library multiple orders of magnitude greater in size than the screening library utilised in the following chapters.

Chapter 5 *In vivo* screening identifies novel drivers in RAS-mutant iCCA

5.1 Introduction

A primary challenge in cancer biology is the discovery of drivers which exist at low frequency in the patient population. The previous chapters have computationally delineated candidate driver genes which occur at low frequency in the input population, and validated a system for delivery of CRISPR/Cas9 editing plasmids into the adult murine liver. In this chapter, the set of predicted drivers is screened *in vivo* alone and in the presence of the Notch Intracellular Domain (NICD) which induces hepatocyte transdifferentiation, or oncogenic RAS^{G12} isoforms, which potentiate the RAS/RAF/MEK/ERK kinase cascade. Both Notch and RAS signalling are dysregulated in iCCA, and how the set of low-frequency predicted drivers may interact with these common aberrations may reveal undiscovered aspects of iCCA biology.

5.2 Hypothesis

Novel low-frequency drivers occur in iCCA and may operate alone or interact with more common molecular aberrations. This question can be explored through *in vivo* screening in the adult murine liver.

5.3 Aims

- The generation of a fully and equally represented SB-CRIPSR plasmid library suitable for unbiased screening *in vivo*.
- To identify driver genes capable of initiating iCCA of hepatocyte origin.
- To identify driver genes cooperating with RAS oncogenes to initiate iCCA of hepatocyte origin.
- The characterisation of screen-derived tumours.

5.4 Generation of a CRISPR/Cas9 screening library

To produce a sgRNA library targeting the candidate driver genes generated in Chapter 3, genes were first searched in the Mouse Genome Informatics database to determine their mouse homologues. This resulted in the dropout of 5 genes for which no mouse homologue exists: KRTAP12-1, encoding a keratin associated protein with putative involvement in hair colour not expressed in the liver[344]; RBMXL1, a primate-specific retrotransposed pseudogene of RBMX also predicted as a driver with 96% sequence conservation and liver expression, but with indeterminate function[345]; SPANXD, a spermatogenesis gene with loss of expression in testicular embryonal carcinoma[346]; ZNF181, a member of the expansive KRAB-ZNF protein family with roles in the suppression of transposable element transcription[347,348]; and C2orf76, a putative protein coding gene with unknown function and limited annotation. This dropout reduced the number of genes for screening to 91.

Two screening libraries targeting the same genes were designed: one for discovery (Lib-A) and the other for validation (Lib-B). Each library contained 473 sgRNAs composed of 273 gene-targeting (3 sgRNAs per gene for 91 drivers), and 200 non-targeting control sgRNAs. Between the libraries the control sequences were the same, but the gene-targeting sequences were different. This was to ensure that genes enriched in Lib-A screening were the result of target gene knockout and not off-target effects: there should be near-perfect concordance between enriched genes in screens performed with Lib-A and Lib-B. sgRNA sequences were obtained

from the Mouse GeCKOv2 library which is designed to limit off-target effects and knock out genes reliably by targeting constitutively expressed exons[287].

To generate plasmid libraries for hydrodynamic tail vain injection, an initial protocol was developed based on those published by the Luo[289] and Zhang[349] labs. This protocol was sequentially optimised to increase yield and library representation (fig 5.4.1; Materials and Methods). sgRNA insert sequences from the resulting libraries were amplified and sequenced on the Illumina MiSeq prior to injection to determine sgRNA representation. The resulting amplicon sequencing validated a pair of libraries with complete (Lib-A) and near-complete (Lib-B) representation, and near-perfect equality as measured by the Gini coefficient of the read count distribution. The Gini index is a metric borrowed from economics where it traditionally measures levels of income inequality - Gini coefficients range from 0 (perfect equality) to 1 (perfect inequality) - but has been used in biology for applications in molecular biology[290], transcriptome analysis[350], and ecology[35] . Here it is used such that a coefficient of 0 would mean that every sgRNA sequence had an identical number of sequencing reads covering it, indicating perfectly equal representation in the plasmid library; whilst a coefficient of 1 would mean every read covered 1 sgRNA sequence, indicative of only one sgRNA sequence present in the library. Here SB-CRISPR-Lib-A had a Gini coefficient of 0.07 and no unrepresented sgRNAs, and SB-CRISPR-Lib-B had a Gini coefficient of 0.09 with 6 unrepresented sgRNAs. The sgRNAs missing from Lib-B were a mix of non-targeting controls (3) and sgRNAs targeting genes (3), however those genes still had 2 sgRNAs covering them.

The design and optimisation of a library generation protocol resulted in a pair of libraries which could be confidently taken forward for *in vivo* screening without the risk of biased library representation influencing the outcome of the screen.

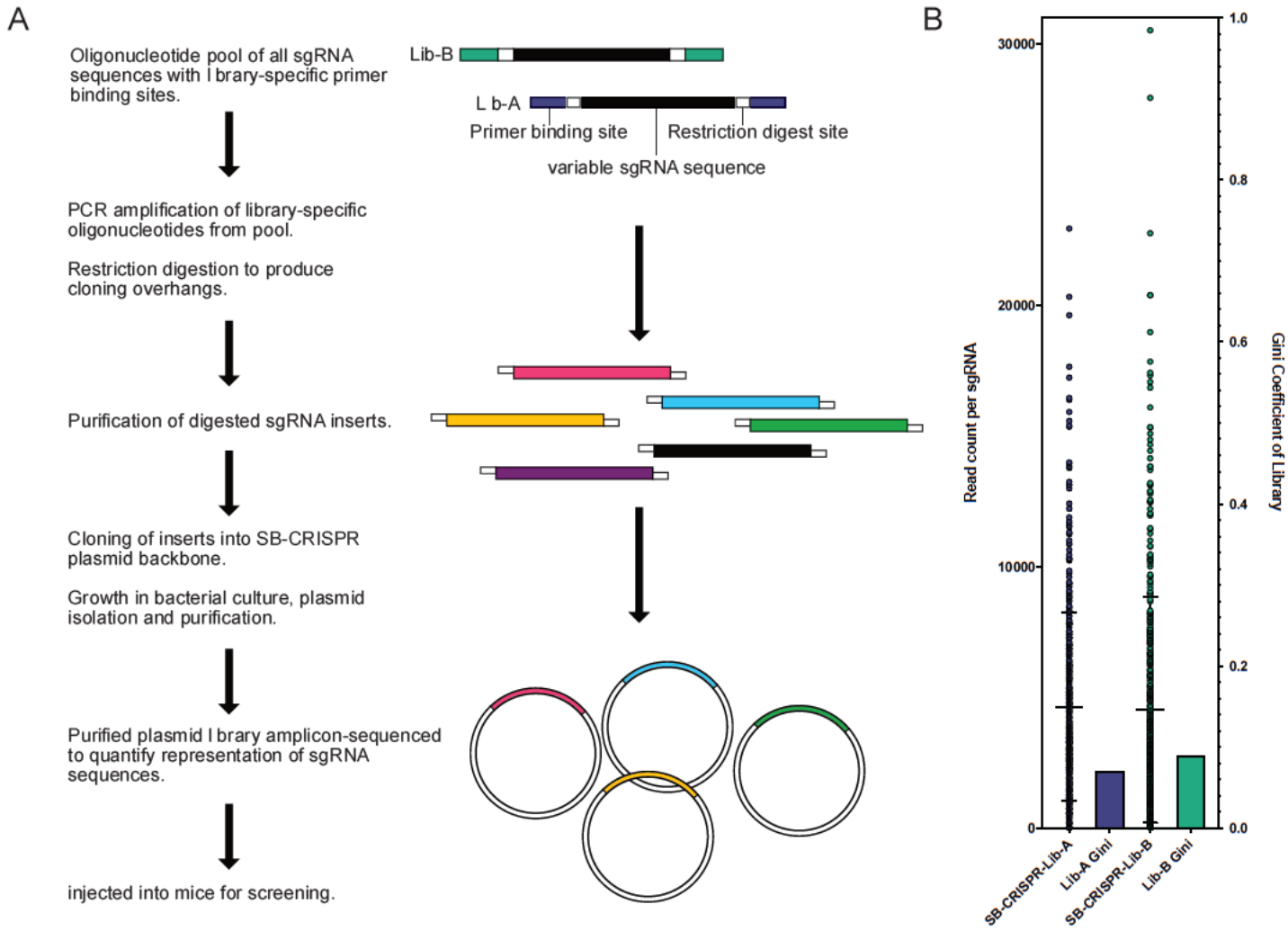
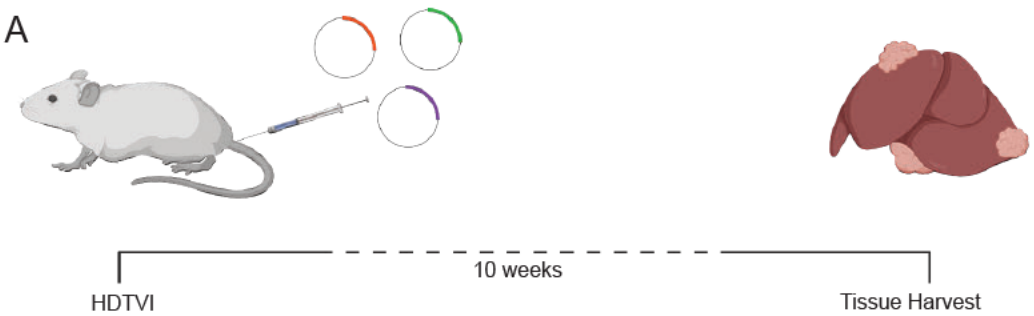
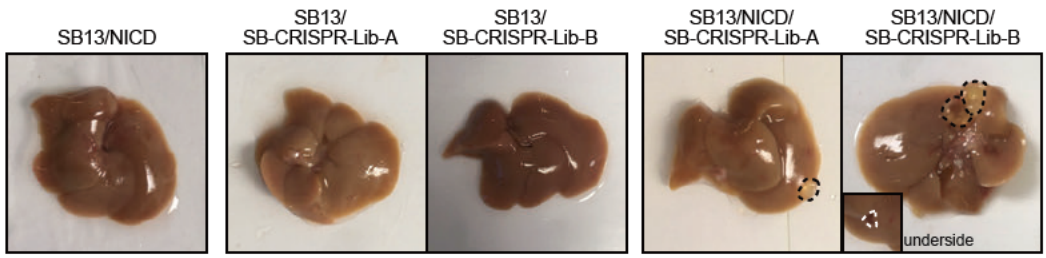

A
Oligonucleotide pool of all sgRNA sequences with l brary-specific primer binding sites.
PCR amplification of library-specific oligonucleotides from pool.
Restriction digestion to produce cloning overhangs.
Purification of digested sgRNA inserts.
Cloning of inserts into SB-CRISPR plasmid backbone.
Growth in bacterial culture, plasmid isolation and purification.
Purified plasmid l brary amplicon-sequenced to quantify representation of sgRNA sequences.
injected into mice for screening.
Lib-B
L b-A
Primer binding site
Restriction digest site
variable sgRNA sequence
B
Read count per sgRNA
Gini Coefficient of Library
30000
20000
10000
0
1.0
0.8
0.6
0.4
0.2
0.0
SB-CRISPR-Lib-A
Lib-A Gini
SB-CRISPR-Lib-B
Lib-B Gini

Figure 5.4.1 Screening library generation and sgRNA sequence representation. A, schematic of how the library is made; libraries are amplified from a pool and cloned into the SB-CRISPR backbone. B, MiSeq amplicon sequencing showed a low Gini index indicative of equal sgRNA sequence representation.

5.5 *In vivo* driver screening generates tumours of hepatic origin

To initially identify combinations of cooperating tumourigenic drivers within the candidate set, the screening library was screened alongside a notch intracellular domain (NICD) expression vector. Notch signalling is a critical determinant of hepatocyte-cholangiocyte/cholangiocarcinoma transdifferentiation in several established models[46,342,352]. However, NICD alone produces no histological change after 10 weeks, and produces cystic cholangiocellular tumours after 20 weeks[99]. Because HDTVI targets hepatocytes and not cholangiocytes, the inclusion of NICD as a transdifferentiation factor was necessary to preferentially screen for drivers involved in iCCA induction, and not HCC induction. NICD, transposase (SB13) and either SB-CRISPR-LibA or SB-CRISPR-LibB were injected into 4-5 week old FVB mice that were subsequently culled 10 weeks post-injection. Mice which received either of the libraries alongside SB13 and NICD developed macroscopically cystic lesions with low penetrance (here defined as the percentage of mice harbouring at least 1 lesion). Whilst those mice that received SB13+library or SB13+NICD did not develop any identifiable lesions after 10 weeks. Due to the cystic nature of the lesions, histology was not possible as tissue collection for DNA and RNA extraction destroyed the structure. This result showed that there are drivers within the screening set which can accelerate the formation of cystic lesions previously described at 20 weeks with NICD alone[99]. Because screening the library without NICD did not result in tumour formation in any mice (n=14) this suggested that the


A
10 weeks
HDTVI
Tissue Harvest


B
SB13/NICD
SB13/
SB-CRISPR-Lib-A
SB13/
SB-CRISPR-Lib-B
SB13/NICD/
SB-CRISPR-Lib-A
SB13/NICD/
SB-CRISPR-Lib-B
underside

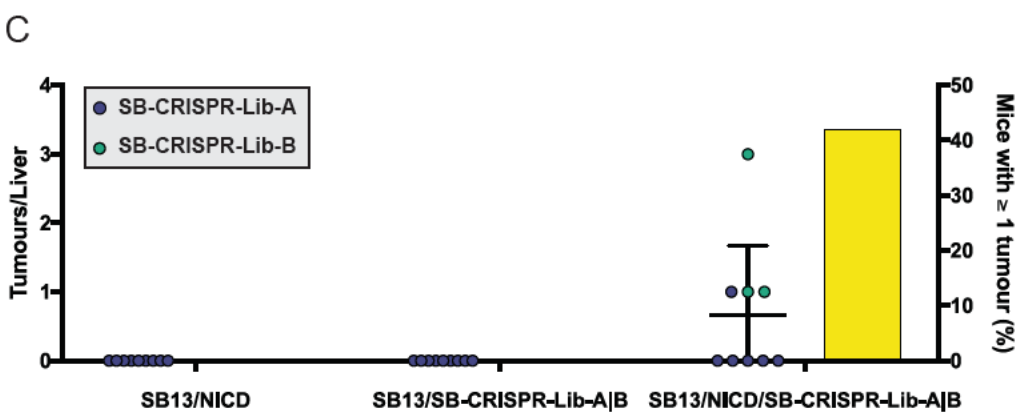

C
SB-CRISPR-Lib-A
SB-CRISPR-Lib-B
Tumours/Liver
Mice with ≥ 1 tumour (%)
0
1
2
3
4
0
10
20
30
40
50
SB13/NICD
SB13/SB-CRISPR-Lib-A|B
SB13/NICD/SB-CRISPR-Lib-A|B

Figure 5.5.1 Experimental design for *in vivo* driver screening and results of NICD/SB-CRISPR-Lib-A A screening. A, mice are injected via hydrodynamic tail vein injection and culled at 10 weeks. B, C, NICD or the screening libraries alone did not produce any tumours, but together resulted in cystic lesion formation at 10 weeks. Only 40% of mice showed this phenotype.

targeted genes could not initiate iCCA of hepatocyte origin without the inclusion of an oncogene.

As discussed in the Introduction, both oncogenic KRAS and oncogenic NRAS are found in iCCA at different frequencies. Whether both RAS oncogenes cooperate with the same or a different set of drivers may reveal underlying common co-dependencies or important differences that could aid in more granular patient stratification. To explore this question, *in vivo* screening of predicted drivers was carried out in the presence of either human $KRAS^{G\ 2D}$ or $NRAS^{G\ 2V}$ expression plasmids.

Following injection of screening libraries, mice injected with either $KRAS^{G\ 2D}$ or $NRAS^{G\ 2V}$ expression plasmids plus either SB-CRISPR-Lib-A or -B harboured large macroscopic tumours consistent with a mass forming (MF) subtype, with phenotypic penetrance comparable between RAS oncogenes. In contrast, mice injected with either of the oncogenic RAS plasmids alone developed neither macroscopic nor microscopic lesions. Whilst differences in tumour burden (number/mouse) between $KRAS^{G\ 2D}$-library and $NRAS^{G\ 2V}$-library mice was not significantly different (P=0.9, Kolmogorov-Smirnov), several $KRAS^{G\ 2D}$-library mice harboured comparatively large numbers of tumours (12, 7, 6) when compared with the $NRAS^{G\ 2V}$-library group (max=3). Additionally, the duration of the experiment was 10 weeks, however all mice in the $KRAS^{G\ 2D}$-library group had to be culled at 8 weeks due to sickness, with several showing ascites. Although not significant at this sample size, it is in support of previous observations regarding the differential oncogenicity of $K/NRAS^{G\ 2}$ isoforms in the liver [93].

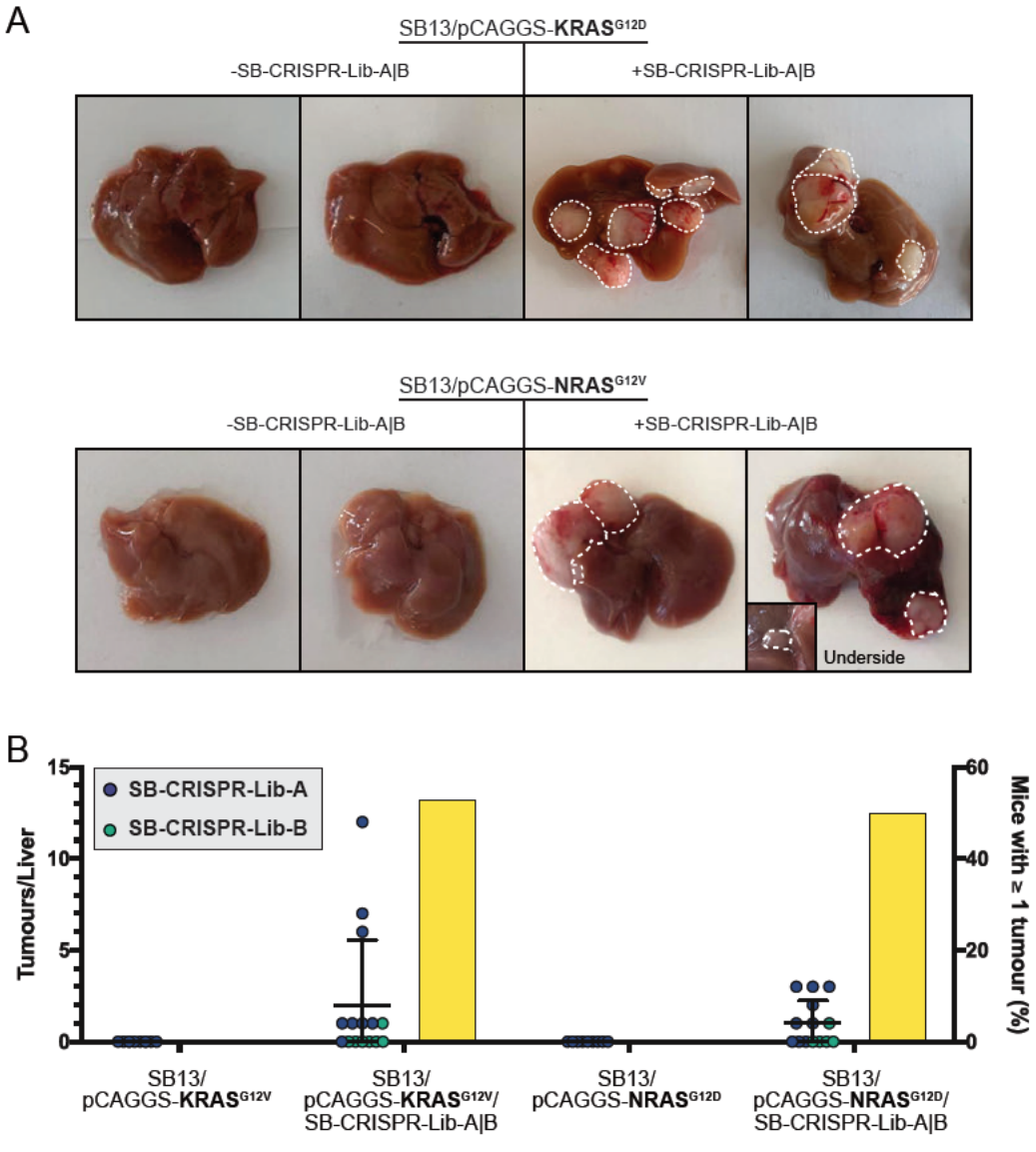
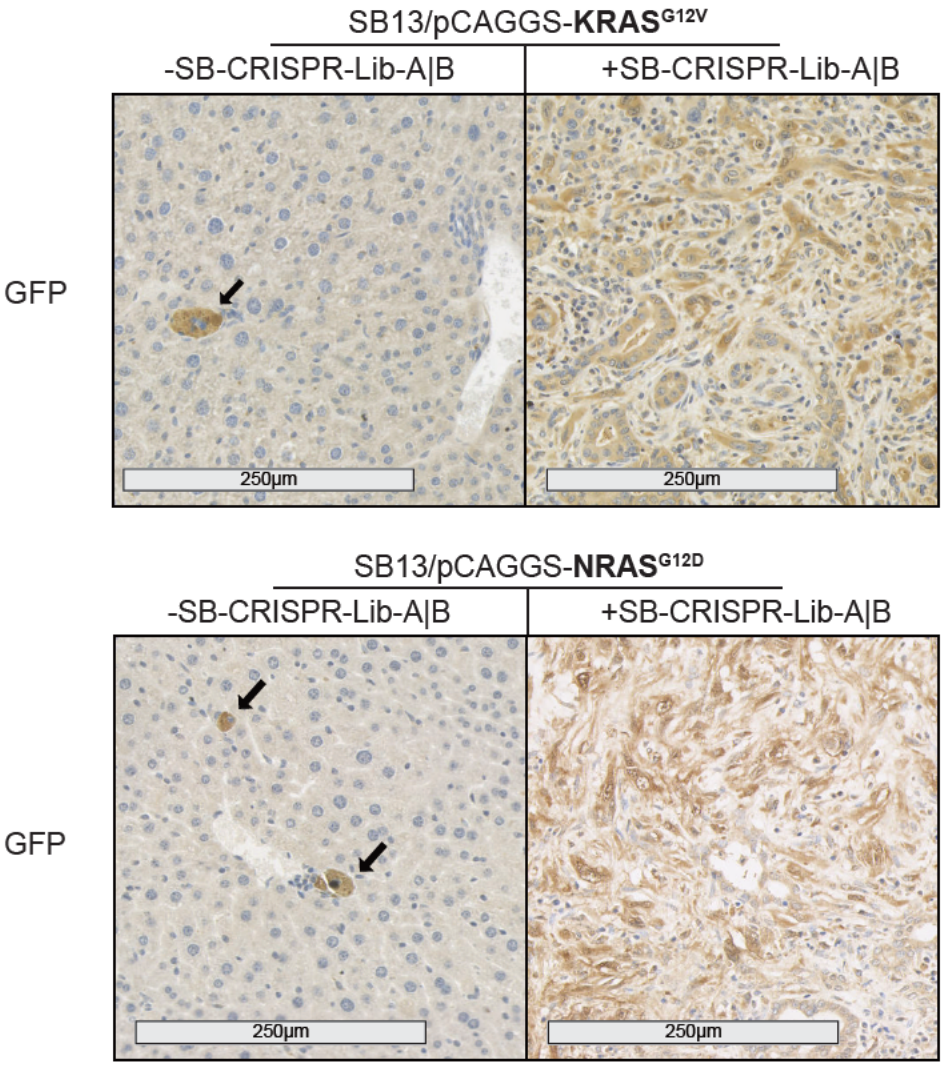

A
SB13/pCAGGS-KRASG12D
-SB-CRISPR-Lib-A|B
+SB-CRISPR-Lib-A|B
SB13/pCAGGS-NRASG12V
-SB-CRISPR-Lib-A|B
+SB-CRISPR-Lib-A|B
Underside
B
SB-CRISPR-Lib-A
SB-CRISPR-Lib-B
Tumours/Liver
Mice with ≥ 1 tumour (%)
0
5
10
15
0
20
40
60
SB13/ pCAGGS-KRASG12V
SB13/ pCAGGS-KRASG12V/ SB-CRISPR-Lib-A|B
SB13/ pCAGGS-NRASG12D
SB13/ pCAGGS-NRASG12D/ SB-CRISPR-Lib-A|B

Figure 5.5.2 Driver prediction libraries synergise with RAS^{G12} oncogenes. A, both library A and B generate tumours when injected alongside oncogenic RAS$^{G\ 2}$ isoforms, resulting in large, solid macroscopic lesions. Oncogenic RAS$^{G\ 2}$ plasmids do not produce tumours alone. B, Both RAS$^{G\ 2}$ oncogenes generated tumours alongside screening libraries at similar frequencies.

As with the SB-CRISPR plasmid, both oncogenic RAS plasmids have expression cassettes flanked by *Sleeping Beauty* LTRs to allow stable genome integration in the presence of SB13 transposase. Both RAS$^{G\ 2}$ expression cassettes are driven from the CAGGS promoter and are transcriptionally linked to GFP via an internal ribosomal entry site. Immunohistochemistry for GFP showed that both RAS$^{G\ 2}$ plasmids had stable tumour integration without expression in the stroma (fig 5.5.3; A). In the livers of mice that had received only the RAS$^{G\ 2}$ plasmid without the screening library, GFP-positive cells were found only very rarely. This is in line with what is known about immune-mediated senescence surveillance and clearance of oncogenic RAS-expressing hepatocytes[243].

Previous studies have reported not to have detected stable integration of the SB-CRISPR plasmid following hydrodynamic delivery to the liver[286]. However, In light of the significantly greater number of editing events seen in Chapter 4 when SB-CRISPR is delivered with SB13 compared to when it is delivered without transposase (fig 4.5.1), whether or not the plasmid was stably integrating in this instance was unclear. To address this, RT-qPCR for Cas9 was carried out using RNA extracted from NRAS$^{G\ 2D}$-library and KRAS$^{G\ 2V}$-library tumours and NICD-library lesions, and compared against mTmG-Cas9 organoids described in Chapter 4 (positive control), and both freshly isolated murine bile duct, and tumours initiated with KRAS$^{G\ 2V}$ and a short hairpin RNA against Trp53 (negative controls). This showed that Cas9 was expressed in NICD-library lesions and in a subset of NRAS$^{G\ 2V}$-library tumours but not all and not in any of the KRAS$^{G\ 2D}$-library tumours (fig 5.5.3; B).

A


SB13/pCAGGS-KRASG12V
-SB-CRISPR-Lib-A|B
+SB-CRISPR-Lib-A|B
GFP
250μm
250μm
SB13/pCAGGS-NRASG12D
-SB-CRISPR-Lib-A|B
+SB-CRISPR-Lib-A|B
GFP
250μm
250μm

B

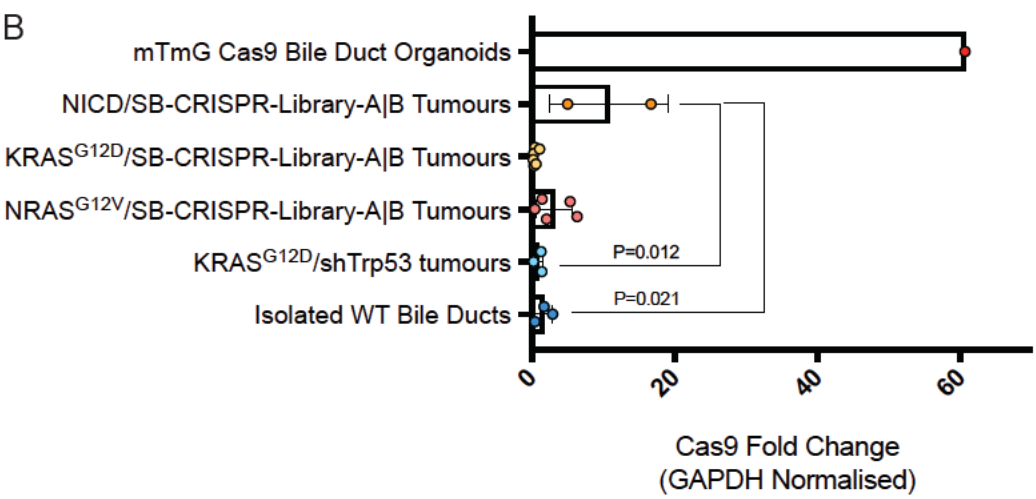

mTmG Cas9 Bile Duct Organoids
NICD/SB-CRISPR-Library-A|B Tumours
KRASG12D/SB-CRISPR-Library-A|B Tumours
NRASG12V/SB-CRISPR-Library-A|B Tumours
KRASG12D/shTrp53 tumours
Isolated WT Bile Ducts
P=0.012
P=0.021
0
20
40
60
Cas9 Fold Change
(GAPDH Normalised)

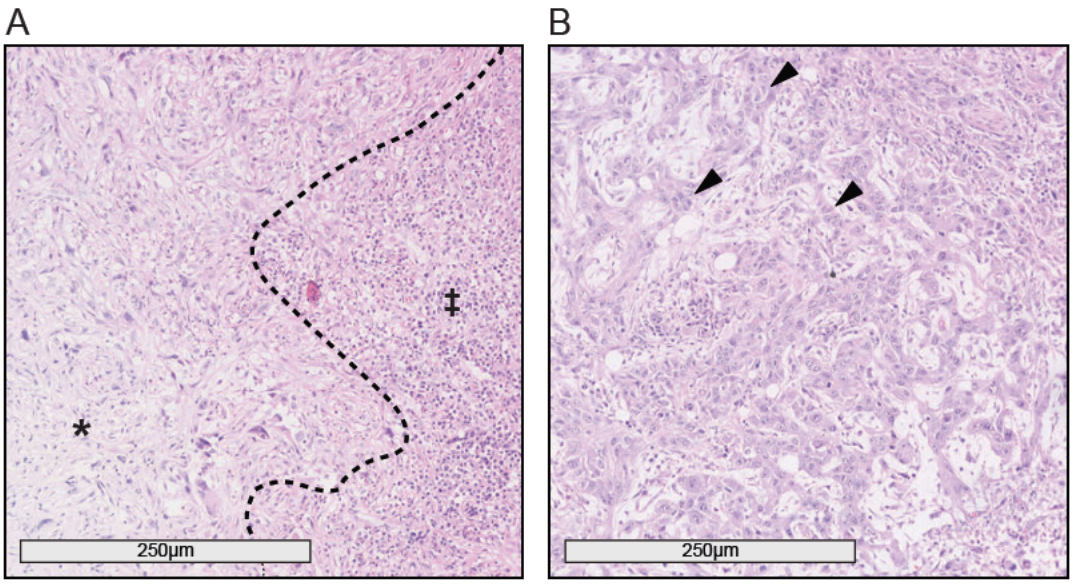
Figure 5.5.3 KRAS and NRAS oncogenes are expressed in tumour cells, but Cas9 expression is significantly decreased. A, in library tumours GFP is localised to the tumour epithelium indicative of oncogenic RAS expression in this compartment. GFP positive cells are rarely seen in the livers of mice that did not receive the screening library. B, Cas9 expression was reduced or undetectable in the screen tumours, showing no significant change compared to the negative controls.

The reason for this heterogeneity is unclear: the integration and expression of the RAS$^{G\ 2}$ expression plasmids demonstrates that the SB13 is delivered and expressed in these cohorts, and the results reported in Chapter 4 further supports this for SB-CRISPR. Possible explanations include the large stromal component of these tumours (discussed below) which may limit the collection of RNA from tumour cells and dilute a Cas9 signal below the limit of detection. Alternatively, cells which have transcriptionally silenced Cas9 could be adaptively selected for during the early stages of tumour evolution. As Cas9 is a foreign entity which elicits an immune response[353 355], its transcriptional silencing may be advantageous in immune evasion. The problem of an anti-Cas9 immune response has been addressed in 2 relevant mouse studies. In one such study[356] the investigators used an adenoviral CRISPR/Cas9 system to edit Pten in the murine liver to model non-alcoholic steatohepatis (NASH). They found anti-Cas9 IgG1/2a antibodies at 14 days post-injection and produced some evidence for a cellular immune response. Nonetheless, they detected CRISPR-mediated Pten mutations and were able to successfully model NASH. In another study, the authors reported T-cell mediated rejection of Cas9-expressing tumour cells in murine cell-line xenograft models. They discovered that following tumour remission several weeks after grafting, there was recurrence of tumours composed of cells with reduced Cas9 expression[357]. This supports the hypothesis that there is an immune-mediated selective pressure for the emergence of Cas9 lowly-expressing cells. Although not explored further here, this is an important consideration when interpreting the output of screening experiments here and elsewhere.

5.6 RAS^{G12}-library tumours recapitulate histological aspects of advanced human iCCA

Human iCCA shows a spectrum of morphological features which has led to multiple histological classification systems[20,358]. Prominent features in classical human iCCA are well-to-moderately differentiated glandular structures surrounded by extensive desmoplastic stroma, whilst more advanced malignancy often presents as poorly differentiated carcinoma lacking well-defined glandular structures [7,359]. In both $KRAS^{G\ 2D}$-library and $NRAS^{G\ 2V}$-library cohorts the tumours showed morphologic heterogeneity within and between tumours. Examination by a clinical histopathologist (Dr Timothy Kendall, Division of Pathology, University of Edinburgh) confirmed that these tumours are iCCA and that they show classical aspects described above, as well as regions of high-grade malignancy and regions of the less commonly observed sarcomatoid morphology. Some tumours also had dense necrotic areas and all showed some degree of immune infiltration as will be discussed. Interestingly, GFP staining (indicative of $RAS^{G\ 2}$ plasmid expression) is seen localised to ductular structures with cuboidal epithelium indicative of the cholangiocyte lineage (fig 5.5.3; A). Because HDTVI targets hepatocytes and not cholangiocytes and these tumours were formed in the absence of NICD, this indicates that $RAS^{G\ 2}$ mutations in combination with targets in the screening library are able to cause transdifferentiation of mature hepatocytes to a differentiated iCCA lineage.

SB13/pCAGGS-**KRAS**^{G12V}/SB-CRISPR-Lib-A|B



SB13/pCAGGS-**NRAS**^{G12D}/SB-CRISPR-Lib-A|B

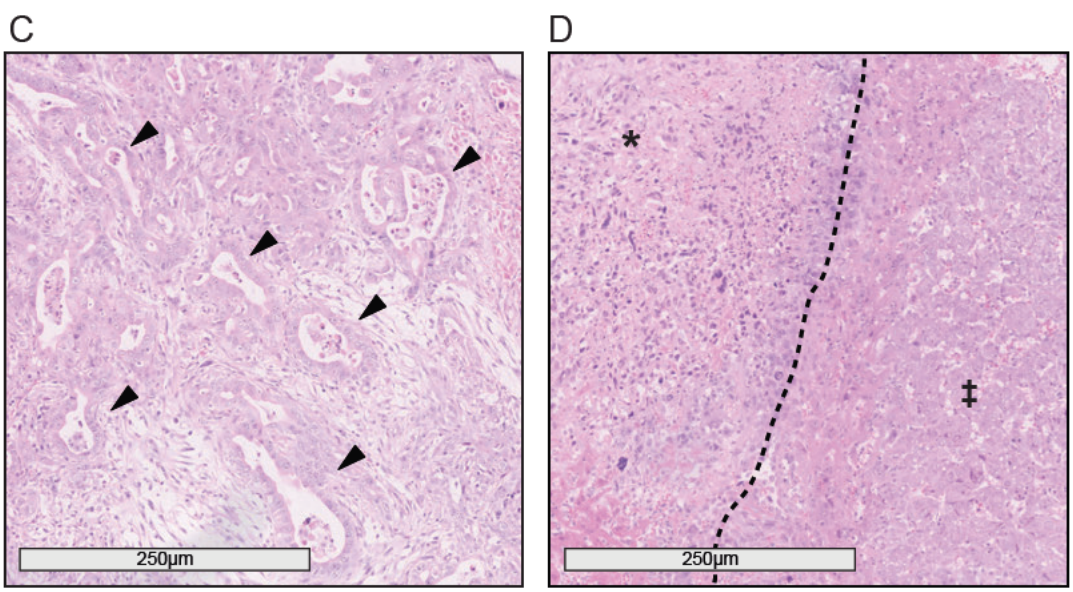


Figure 5.6.1 RAS^{G12} screen tumours show variable morphologies. A-B, KRAS$^{G\ 2D}$ library tumours. Arrowheads highlight cuboidal epithelium with loss of definitive ductular structures; asterisk denoted desmoplastic stroma, and double-dagger denotes necrotic region. C-D, arrowheads denote cuboidal epithelium organised into ductular structures. Asterisk indicates regions classified as ‘sarcomatoid’ and double-dagger denotes region classified as ‘high-grade’.

To further investigate the identity of tumour cells, sections were stained with a broad-spectrum cytokeratin antibody (PanCK) and the cholangiocyte marker cytokeratin 19 (CK19) which has value in differentiating iCCA from HCC[360,36] . All tumours showed pervasive cytokeratin expression indicative of carcinoma and this also served to highlight necrotic and stromal regions for example necrotic cores can be seen in figure 5.6.2 A (densely packed cells, blue regions in macro images, and on the right of insets). The cholangiocyte marker CK19 has tightly restricted expression in normal liver where it is confined to the bile duct epithelium and not expressed in any other cell type. Both KRAS$^{G\ 2D}$ and NRAS$^{G\ 2V}$ tumours showed extensive CK19 expression in cells of variable morphologies. Well-differentiated ductular structures with CK19+ cuboidal epithelium were readily observed in all tumours, with a subset also showing CK19+ cuboidal epithelium not arranged in glands. Remarkably, both RAS$^{G\ 2}$ tumour cohorts also had cases with CK19+ spindle cells recapitulating the morphology of rare intrahepatic sarcomatous tumours[362,363].

A predominant feature of iCCA is its fibro-desmoplastic stroma, consisting of dense extracellular matrix (ECM) composed of ECM molecules not present in normal liver[364]. Cellular contributions to the stroma include a dynamic repertoire of immune cells, and cancer associated fibroblasts (CAFs) which are largely responsible for establishing the tumour ECM[365].

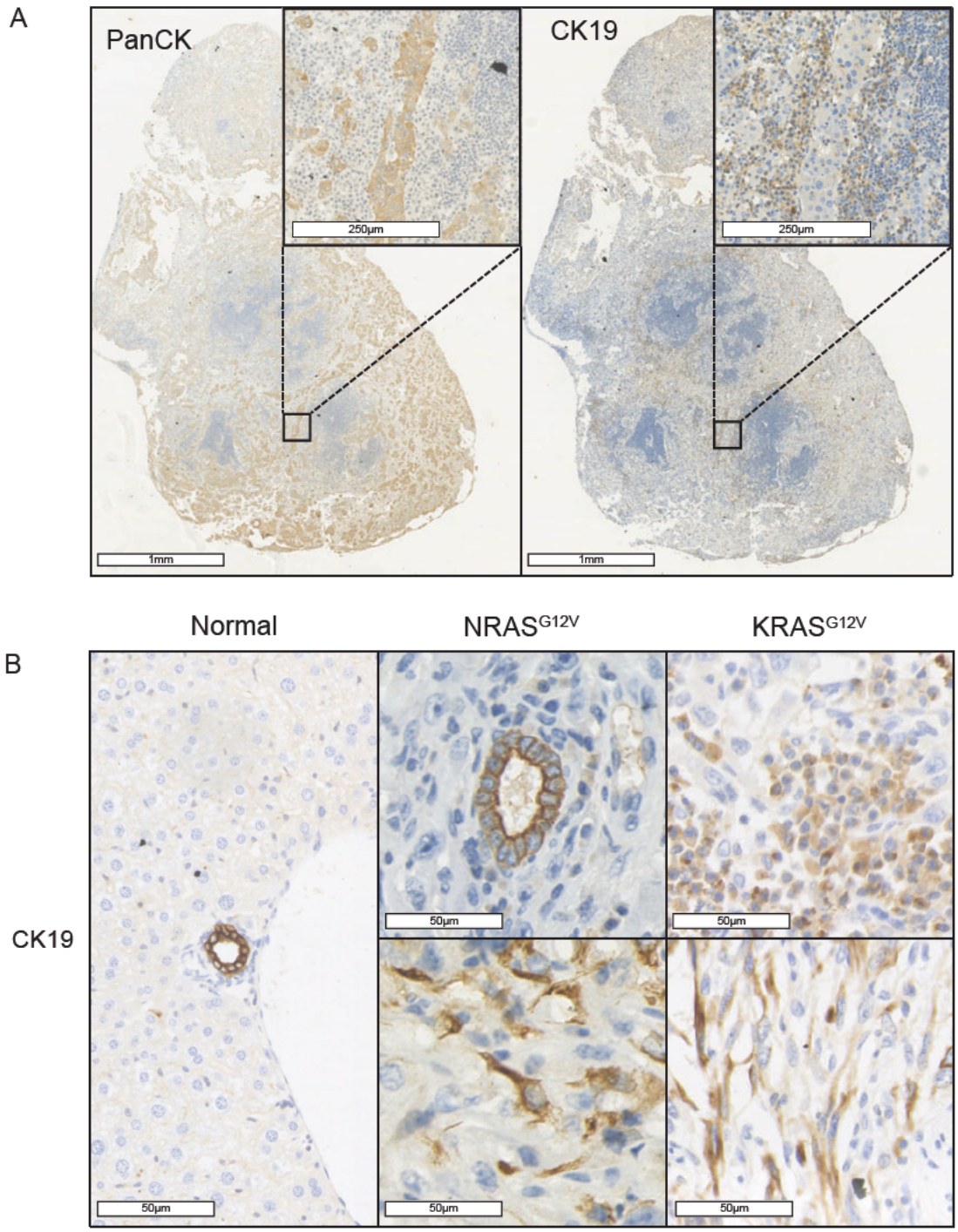

A
PanCK
CK19
250μm
250μm
1mm
1mm
Normal
NRAS^G12V
KRAS^G12V
B
CK19
50μm
50μm
50μm
50μm
50μm

Figure 5.6.2 Lineage identity of RAS^{G12} screen tumours. A, PanCK stains cytokeratin, a marker for tumour epithelium and CK19 highlights the cholangiocyte identity of much of the epithelial compartment. Comparison of both images demonstrates that some regions lose CK19 indicating loss of biliary cell identity. B, in normal liver CK19 is tightly restricted to the bile duct epithelium, but in RAS$^{G\ 2}$ tumours CK19 staining cells are found in several morphologies; ductular, cuboidal, and spindle-like.

To explore the acellular composition of the stroma, staining for collagen 1, collagen 6A6, and pan-fibronectin was carried out. Collagens, generally speaking, are important stromal components which modulate cancer cell biology and prognosis[366]. The choice of Col1 and Col6A6 here is based on the formers previous association with cholangiocarcinoma as a biomarker for fibrosis[367] due to its secretion by cancer associated fibroblasts (CAFs); and the latter due to its strong differential expression between normal and NRAS$^{G\ 2V}$ tumours in this study (data not shown). Similar to Col1, fibronectin is also secreted by reactive CAFs which have subsequently been shown to align fibronectin to orchestrate directional cancer cell migration[368].

Col1 was extensively present in the stroma of all KRAS$^{G\ 2D}$ and NRAS$^{G\ 2V}$ tumours. When compared to the NICD/myrAKT model, Col1 staining serves to highlight the comparatively dense stroma seen in screen tumours and suggests that this model may be a useful tool for studying stroma in iCCA. Col6A6 was present in normal liver with peri-portal staining which is extracellular and localised to the basement membrane surrounding the ducts. Curiously, in both RAS$^{G\ 2}$ and NICD/myrAKT tumours, staining was seen in both the intracellular and extracellular compartments. Col6A6 is one of 6 Col6 chains which make up the Col6 protein[369]. Col6 has a been shown to activate canonical Wnt signalling[370], as well as promote epithelial-mesenchymal transition and tumour inflammation[37 373] all important processes linked to iCCA. Importantly, *in vitro* and *in vivo* evidence suggests that Col6 expression is increased in response to cisplatin and that this mediates cisplatin resistance in ovary and breast cancer[374,375]. Whist these studies focused on subtypes

of Col6 other than Col6A6, this is still a relevant finding for iCCA which is treated with cisplatin as standard and often develops resistance; further study of Col6A6 in iCCA is warranted. Fibronectin was found to be weakly expressed in the cytoplasm of normal hepatocytes, but not biliary cells, and this is in line with previous studies assessing fibronectin expression in the normal liver[376]. NICD/myrAKT, and NRAS$^{G\ 2V}$ tumours had strong stromal staining likely as a product of CAFs. Unexpectedly, KRAS$^{G\ 2D}$ tumours did not have strong stromal fibronectin staining, but occasionally had intracellular staining in the cytoplasm of ductular structures. Whilst cytoplasmic accumulation of fibronectin has been shown in CCA[377] and in thyroid carcinoma[378], further validation of this finding is necessary to rule out non-specific staining. In line with Col1 and fibronectin being prominent acellular components of the stroma, α-smooth-muscle actin (αSMA) positive CAFs were also a pervasive feature in both screen tumours and the NICD/myrAKT control tumours. αSMA expressing cells were not present in the normal liver: hepatic fibroblasts and CAFs transdifferentiate from quiescent hepatic stellate cells in response to injury and acquire αSMA expression only when activated[379,380]. Infiltrating macrophages and neutrophils were observed in both RAS$^{G\ 2}$ tumours and NICD/myrAKT tumours, however regulatory T-cells were seen only in RAS$^{G\ 2}$ groups, and not in contemporaneously stained NICD/myrAKT tumours.

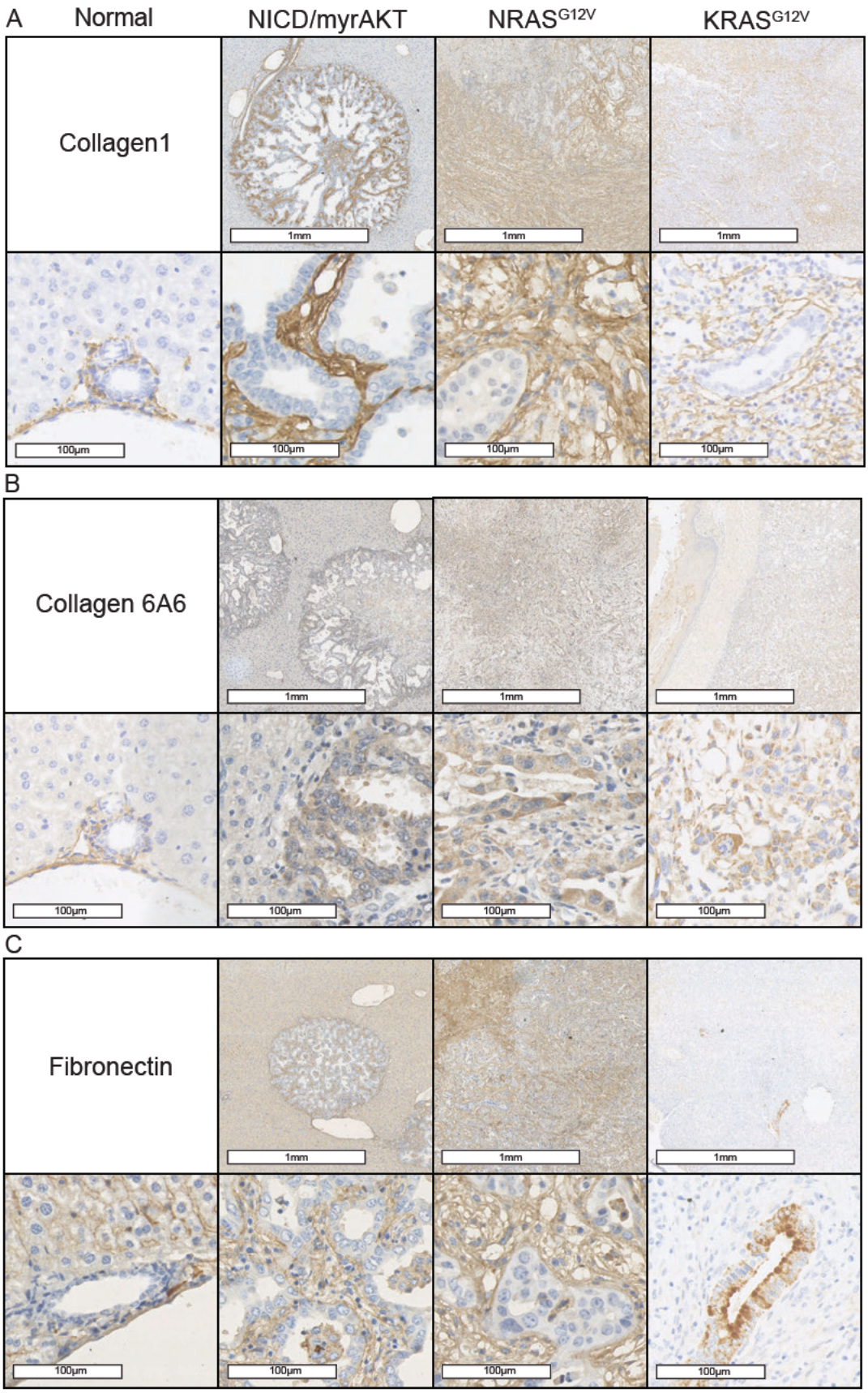

A
Normal
NICD/myrAKT
NRASG12V
KRASG12V
Collagen1
1mm
100μm
B
Collagen 6A6
1mm
100μm
C
Fibronectin
1mm
100μm

Figure 5.6.3 Extracellular matrix components. All tumours stained positive for collagen 1 (A), B, Collagen 6A6 showed basement membrane localisation in normal liver but expression in the cellular and stromal compartments in tumours. C, fibronectin showed the expected localisation in all groups except the $KRAS^{G\ 2D}$ screen group, where it was seen localised to well differentiated ductular structures.

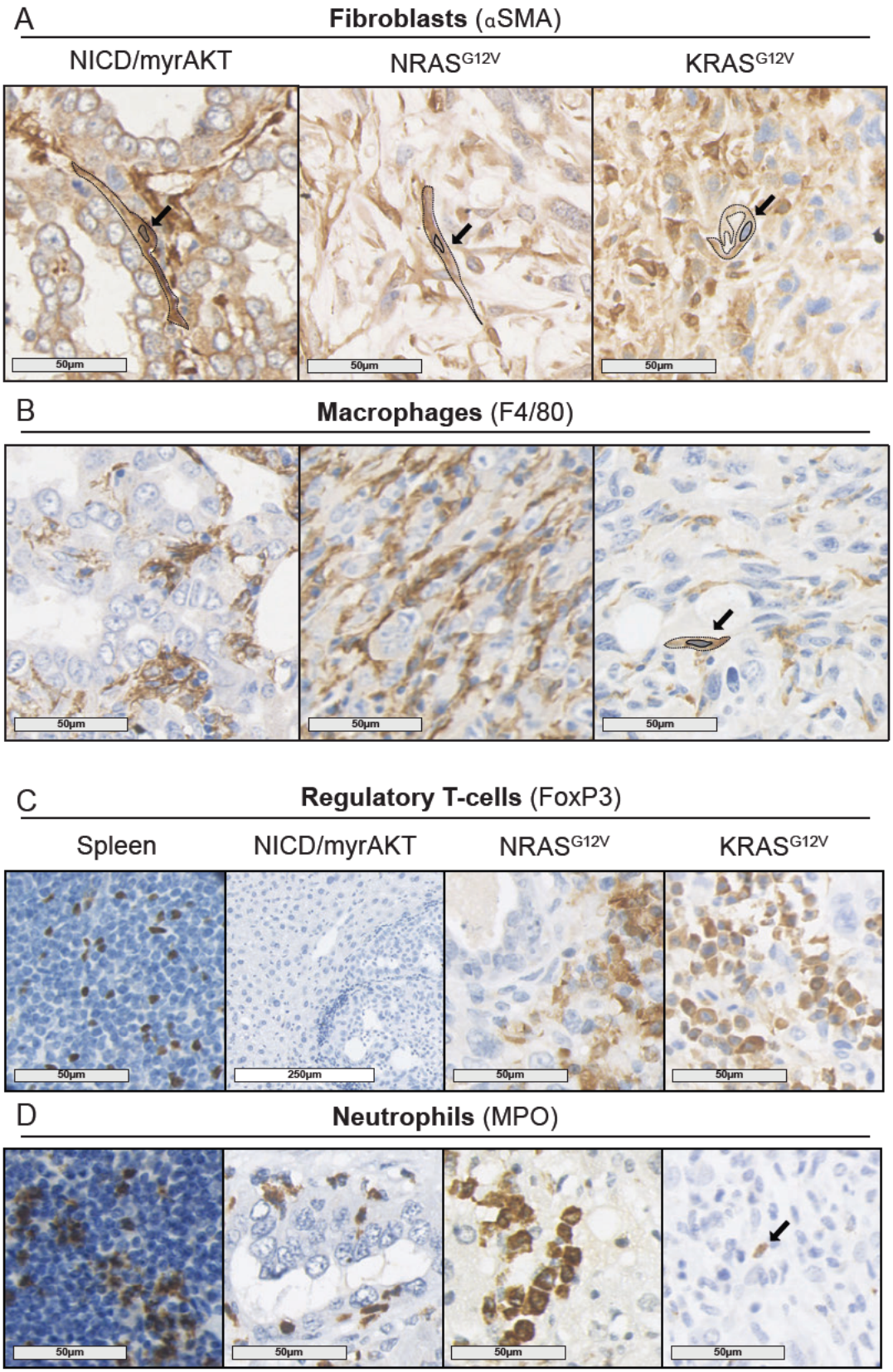

A
Fibroblasts (αSMA)
NICD/myrAKT
NRASG12V
KRASG12V
50μm
50μm
50μm
B
Macrophages (F4/80)
50μm
50μm
50μm
C
Regulatory T-cells (FoxP3)
Spleen
NICD/myrAKT
NRASG12V
KRASG12V
50μm
250μm
50μm
50μm
D
Neutrophils (MPO)
50μm
50μm
50μm
50μm

Figure 5.6.4 Immune infiltrate in tumours. All tumours showed infiltration of cancer associated fibroblasts (A), tumour associated macrophages (B), and neutrophils (D), however only $RAS^{G\ 2}$ screen tumours showed regulatory T-cell infiltration (C). Spleen was used as a positive control.

5.7 Exome sequencing of RAS^{G12}-library tumours identifies novel drivers of intrahepatic cholangiocarcinoma

Exome sequencing was used to determine which CRISPR-mediated mutations generated iCCA in conjunction with KRAS and NRAS oncogenes. 16 $KRAS^{G\ 2D}$-library tumours (including 2 pancreatic lesions, thought to be invasions from the liver), and 10 $NRAS^{G\ 2V}$-library tumours were sequenced alongside one normal liver control from an uninjected mouse. All indels within 50bp of a sgRNA binding site were visually inspected to determine overlap with sgRNA target sites. Virtually all indels had start or end positions approximately 3bp upstream of the Cas9 protospacer adjacent motif (PAM) sequence which strongly indicates that they are the consequence of CRISPR/Cas9 editing and not spurious.

A potential risk of inducing so many CRISPR/Cas9 DSBs simultaneously in individual cells is that they may induce larger structural rearrangements through repair of discordant breakpoints, or excision of intervening sequence between sgRNAs as demonstrated in the mTmG model in Chapter 4. Previous studies that used the SB-CRISPR plasmid for screening a smaller number of genes in the liver reported no such rearrangements, however, due to the comparatively larger set of sgRNAs used this was checked here. Structural variant analysis on the $NRAS^{G\ 2V}$-library tumour exomes confirmed that, whilst not extensive, 4 high-confidence structural rearrangements mapping to CRISPR/Cas9-induced breakpoints were observed in 3 samples. In 2 such cases, inversion of a segment between a Trp53 sgRNA binding site and either a Nf2 or Rnf43 sgRNA binding site was observed

leading to the disruption of target exons in both cases. In another tumour, a small intragenic deletion between 2 Plk2 sgRNA binding sites was seen which served to disrupt only this gene. And lastly, one case harboured a 32kb duplication of a segment spanning part of Nf2 exon 7, with the end of the duplicated sequence mapping to a Nf2 sgRNA target site. Although uncommon, duplications due to CRISPR/Cas9 have been described, and these tended to segregate with deletions or inversions[38] ; this seems to be the case here where the Trp53-Nf2 inversion described above was found in the same tumour as the Nf2 exon 7 duplication.

A minimum of 2 drivers were observed per tumour across groups, with a whole-number median of 5 drivers pre tumour seen in the $NRAS^{G\ 2V}$ group, and 6 (actual 5.5) in the $KRAS^{G\ 2D}$ group. Interestingly, tumours from the same mouse did not cluster together except in the 2 mice with pancreatic invasions. In these mice, pancreatic tumours clustered with 2-3 other samples from the same mouse. As library uptake is random, this suggests that these tumour clusters may be descendants of a common clone/tumour. Nuclear Corepressor 1 (Ncor1) was edited in the pancreatic metastasis of mouse 986418 and was not present in any of the other 986418 tumours that it clustered with. Additionally, Ncor1 was detected in a tumour of the NRAS group which uniquely showed diaphragmatic invasion. Ncor1 expression is downregulated in HCC and breast cancer and its loss in experimental systems causes invasion and metastasis[382], which supports a role for it here as a causative mutation for this phenotype.

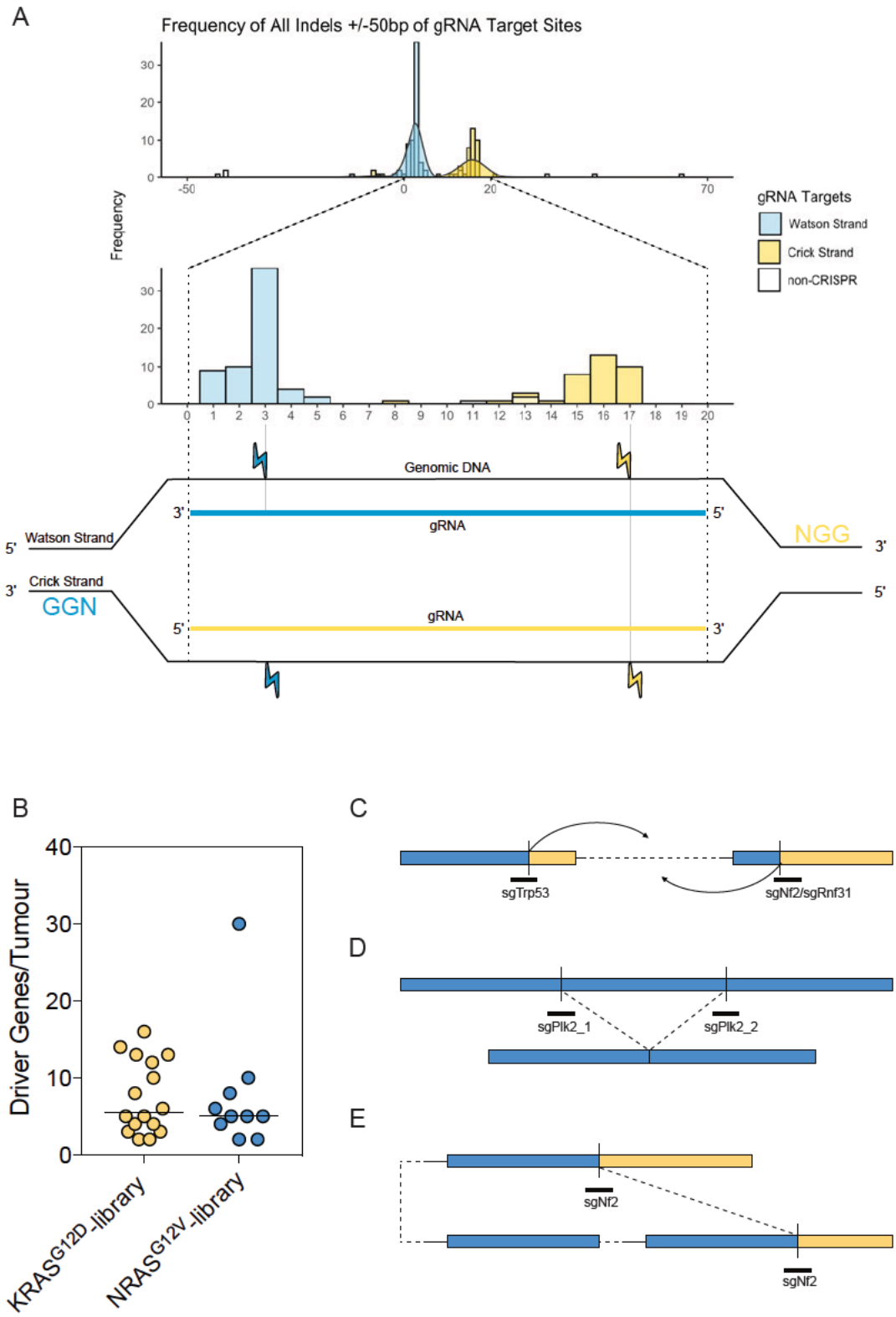

A
Frequency of All Indels +/-50bp of gRNA Target Sites
Frequency
gRNA Targets
Watson Strand
Crick Strand
non-CRISPR
Genomic DNA
5'
Watson Strand
3'
gRNA
5'
NGG
3'
3'
Crick Strand
GGN
5'
gRNA
3'
5'
B
Driver Genes/Tumour
KRASG12D-library
NRASG12V-library
C
sgTrp53
sgNf2/sgRnf31
D
sgPlk2_1
sgPlk2_2
E
sgNf2
sgNf2

Figure 5.7.1 CRISPR/Cas9 editing outcomes. A, indels within 50bp of sgRNA spacer target sites localise 2-4bp proximal to PAM sequences indicative of CRISPR/Cas9 mediated editing. B, Tumours in NRAS and KRAS oncogene screens harboured on average the same number of CRISPR/Cas9-induced mutations (~5). C, schematic of structural variations likely caused by Cas9 editing, with breakpoints localising to sgRNA target sequences (black bars).

Across both RAS$^{G\ 2}$-library screens a total of 66 of the 91 predicted drivers targeted by the screening library were mutated, and 27 of these were shared between both screens. Notably, Trp53 was mutated in all but 2 samples which reinforces the importance of Trp53 loss in the escape of Ras-induced senescence in hepatocytes. Despite TP53 not significantly co-occurring with RAS mutations in the human data this result shows that other genes capable of escaping RAS-induced senescence at least in hepatocytes were not present amongst predicted drivers. This disparity is perhaps explainable by the fact that in this study iCCA is of hepatocyte origin so necessarily requires Trp53 loss to escape senescence.

Broad functional profiling of shared drivers uncovered significantly enriched biological processes including numerous developmental and neurodevelopmental processes, morphogenetic processes, and cell mobility and motility. The overlap between neurodevelopmental and cancer genes is not novel: genes with canonical involvement in neurodevelopment due to their role in cell signalling (Nf2, K/Nras, Cacna1c), cell motility (Plxnb2, Fn1, Thbs1, Ablim1), and cell cycle (Rb1, Plk2) were found mutated in this screen. Some of these have known roles in cancer but some have not previously been described.

At a regulatory level, 19 of the 27 shared drivers had predicted binding sites for the Zic1 transcription factor. Zic1 is significantly decreased in gastric and colorectal cancer through promoter hypermethylation, and associated with decreased overall survival[383 386]. Additionally, Trp53 and Nf2 were both found to interact with miR-138-5p which has tumour suppressive roles in cancer[387], and has previously been linked to regulation of the Hippo pathway through modulation of YAP[388]. Although

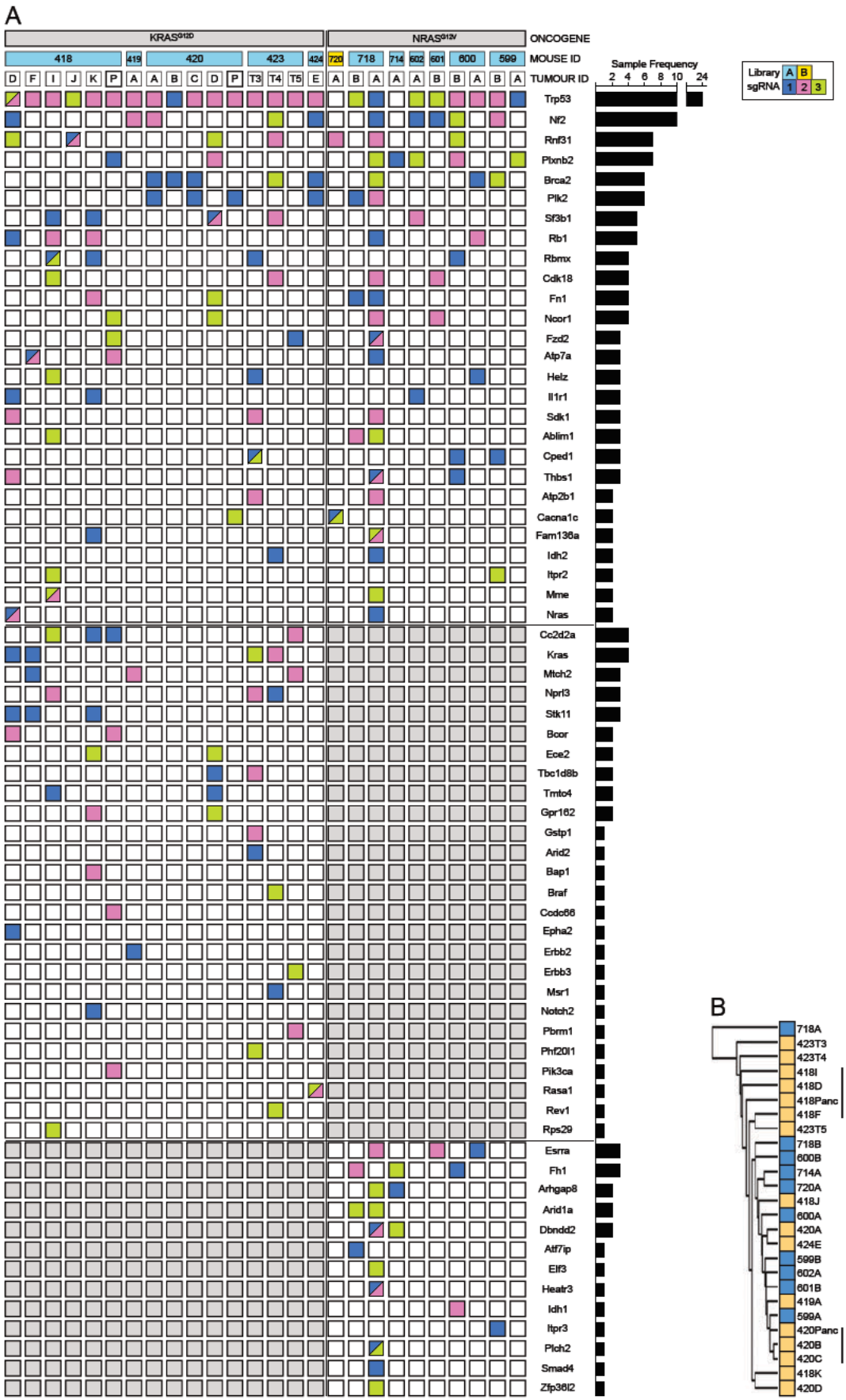
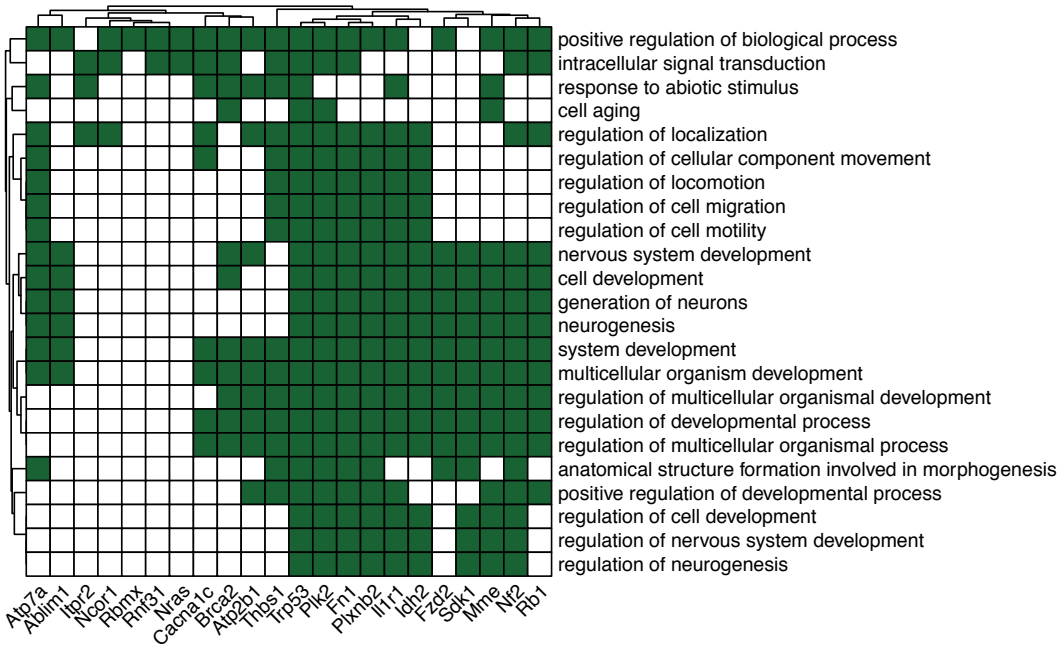
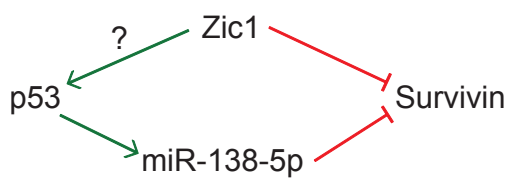

A
KRAS G12D
NRAS G12V
ONCOGENE
MOUSE ID
TUMOUR ID
Sample Frequency
2 4 6 8 10 24
Library A B
sgRNA 1 2 3
Trp53
Nf2
Rnf31
Plxnb2
Brca2
Plk2
Sf3b1
Rb1
Rbmx
Cdk18
Fn1
Ncor1
Fzd2
Atp7a
Helz
Il1r1
Sdk1
Ablim1
Cped1
Thbs1
Atp2b1
Cacna1c
Fam136a
Idh2
Itpr2
Mme
Nras
Cc2d2a
Kras
Mtch2
Nprl3
Stk11
Bcor
Ece2
Tbc1d8b
Tmto4
Gpr162
Gstp1
Arid2
Bap1
Braf
Codc66
Epha2
Erbb2
Erbb3
Msr1
Notch2
Pbrm1
Phf20l1
Pik3ca
Rasa1
Rev1
Rps29
Esrra
Fh1
Arhgap8
Arid1a
Dbndd2
Atf7ip
Elf3
Heatr3
Idh1
Itpr3
Plch2
Smad4
Zfp36l2
B
718A
423T3
423T4
418I
418D
418Panc
418F
423T5
718B
600B
714A
720A
418J
600A
420A
424E
599B
602A
601B
419A
599A
420Panc
420B
420C
418K
420D

Figure 5.7.2 Common and unique drivers between KRAS^{G12D} and NRAS^{G12V} screens. A, Exome-seq results for a subset of tumours showing which driver genes were mutated and with which sgRNA. Trp53 is only rarely targeted by sgRNA-1 which was underrepresented in the screening library when compared to sgRNAs 2 and 3. Notably NRAS tumour 718-A harboured a large number of mutations and this is possibly a result of 718-A being having multiple cells of origin as discussed above. 27 drivers were shared between screens, with Trp53 present in all samples except one across all screens. B, Based on driver gene profiles, KRAS$^{G\ 2D}$ and NRAS$^{G\ 2V}$ screen tumours do not cluster together, nor do tumours from the same liver, with the exception of tumours from mouse 986418 and 986420, these mice also show pancreatic invasions indicative of tumours originating from the same clone.

A


positive regulation of biological process
intracellular signal transduction
response to abiotic stimulus
cell aging
regulation of localization
regulation of cellular component movement
regulation of locomotion
regulation of cell migration
regulation of cell motility
nervous system development
cell development
generation of neurons
neurogenesis
system development
multicellular organism development
regulation of multicellular organismal development
regulation of developmental process
regulation of multicellular organismal process
anatomical structure formation involved in morphogenesis
positive regulation of developmental process
regulation of cell development
regulation of nervous system development
regulation of neurogenesis
Atp7a
Ablim1
Itpr2
Ncor1
Rbmx
Rnf31
Nras
Cacna1c
Brca2
Atp2b1
Thbs1
Trp53
Plk2
Fn1
Plxnb2
Il1r1
Idh2
Fzd2
Sdk1
Mme
Nf2
Rb1

B


Zic1
?
p53
Survivin
miR-138-5p

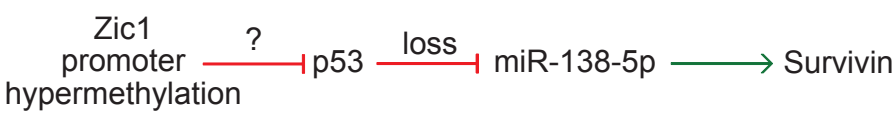

Zic1
promoter
hypermethylation
?
p53
loss
miR-138-5p
Survivin

Figure 5.7.3 Biological process enrichment in shared drivers and proposed Survivin regulatory module. A, Gene Ontology Biological Process terms enriched amongst the shared driver set mainly involved developmental processes and regulation of cell migration. B, the transcription factor Zic1 and microRNA miR-138-5p had proposed binding sites for/interactions with p53. As both also regulate Survivin, which is upregulated in iCCA, a proposed mechanism of whereby Survivin may be upregulated in response to mutational loss of p53 or promoter hypermethylation of Zic1. This depends on whether or not Zic1 ordinarily acts to induce p53 expression.

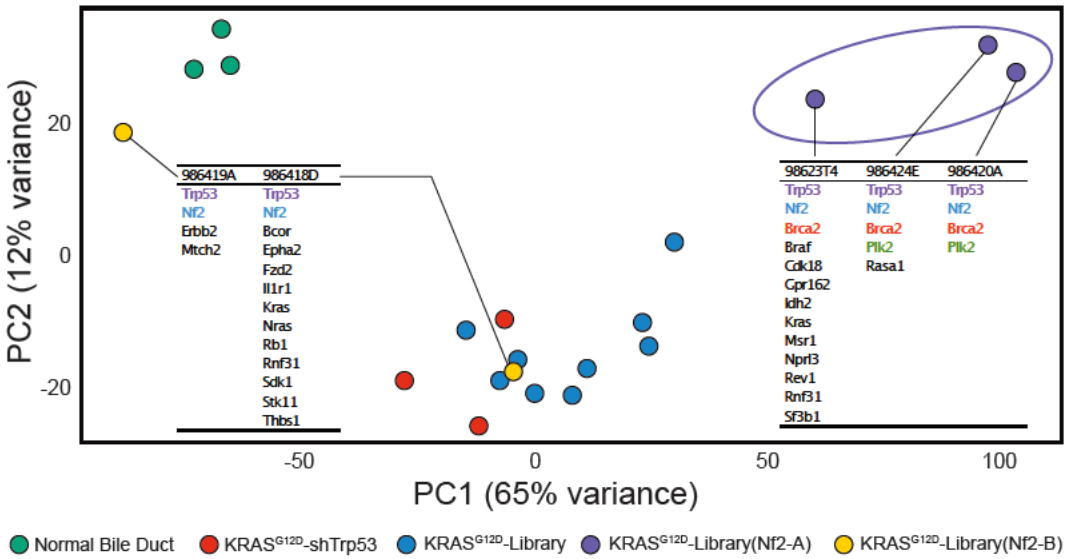
somewhat tangential, it's interesting to note that Zic1 and miR-138-5p both interact with Trp53, and both are negative regulators of Survivin[389,390]. Survivin a negative regulator of apoptosis is expressed in human cholangiocarcinoma in response to the chemotherapeutic agent doxorubicin[39] and is correlated with poor prognosis[392]. Survivin is also differentially upregulated in all tumours in this study as assessed by transcriptome analysis (see below). A 2016 study found that regulation of p53 and miR-138 was unidirectional in human lung cancer cells, with p53 loss resulting in miR-138 downregulation. This hints at a potential mechanism whereby p53 loss may lead to Survivin expression via downregulation of miR-138-5p, priming cells for doxorubicin resistance. If Zic1 expression is similarly decreased in iCCA as it is in gastric and colorectal cancer, this may be an additional route to Survivin expression without mutation of TP53, contingent on Zic1 acting as a positive regulator of TP53 expression.

To further characterise the screen tumours, RNA sequencing (RNA-seq) was used to determine differentially expressed genes when compared to normal bile duct or KRAS$^{G\ 2D}$-shTrp53 tumours. As expected, normal bile duct and KRAS$^{G\ 2D}$-shTrp53 tumours clustered together on principal component analysis, whilst KRAS$^{G\ 2D}$-library tumours did not cluster together, and instead were distributed across the plot. A loose cluster could be seen in the top right of the plot and these 3 samples all had Nf2 mutations. Nf2 was the most frequently mutated gene across both RAS$^{G\ 2}$ screens and splitting the samples based on Nf2 mutation status separated the KRAS$^{G\ 2D}$-library group into 3. Nf2-A consisted of screen tumours harbouring Nf2 mutations which clustered together in the top right corner, Nf2-B consisted of

screen tumours with Nf2 mutations which did not cluster together, and the remaining group consisted of screen tumours without Nf2 mutation. Nf2-A samples were all from different mice and uniquely shared Brca2 mutations when compared to Nf2-B. The 2 samples that clustered closest in the top right additionally shared Plk2 mutations. Indeed, one such tumour harboured only 4 CRISPR-mutated genes: Trp53, Nf2, Brca2, and Plk2.

Differential gene expression analysis between the Nf2-A screen tumours and those without Nf2 mutations identified a large set of differentially downregulated genes, but few upregulated genes (Appendix 2). Enrichment analysis for biological processes showed only weak enrichment for cAMP signalling amongst the 229 upregulated genes, however the 1312 differentially downregulated genes segregated into 741 process annotations. A complete list of enriched processes, cellular compartments, molecular functions, KEGG and REACTOME pathways, and transcription factor binding sites can be found in Appendix 3. As a means of highlighting disease-relevant biological processes, the set of enriched biological processes was filtered based on the following terms ('*' denotes any combination of letters): "cancer", "liver", "bile", "apopto*", "cycle", "signal*", "matrix", "development", "immun*", and "*abol*". Enriched terms (limited to a maximum of 30) are displayed in Fig 5.7.3, A.. Most striking is the strong statistical enrichment of terms related to metabolism, suggesting extensive metabolic rearrangements in these tumours. Processes linked to regulation of apoptosis, immune response, and ECM composition were also enriched, supporting the observed immune infiltrate

A



Normal Bile Duct KRAS^{G12D}-shTrp53 KRAS^{G12D}-Library KRAS^{G12D}-Library(Nf2-A) KRAS^{G12D}-Library(Nf2-B)

B

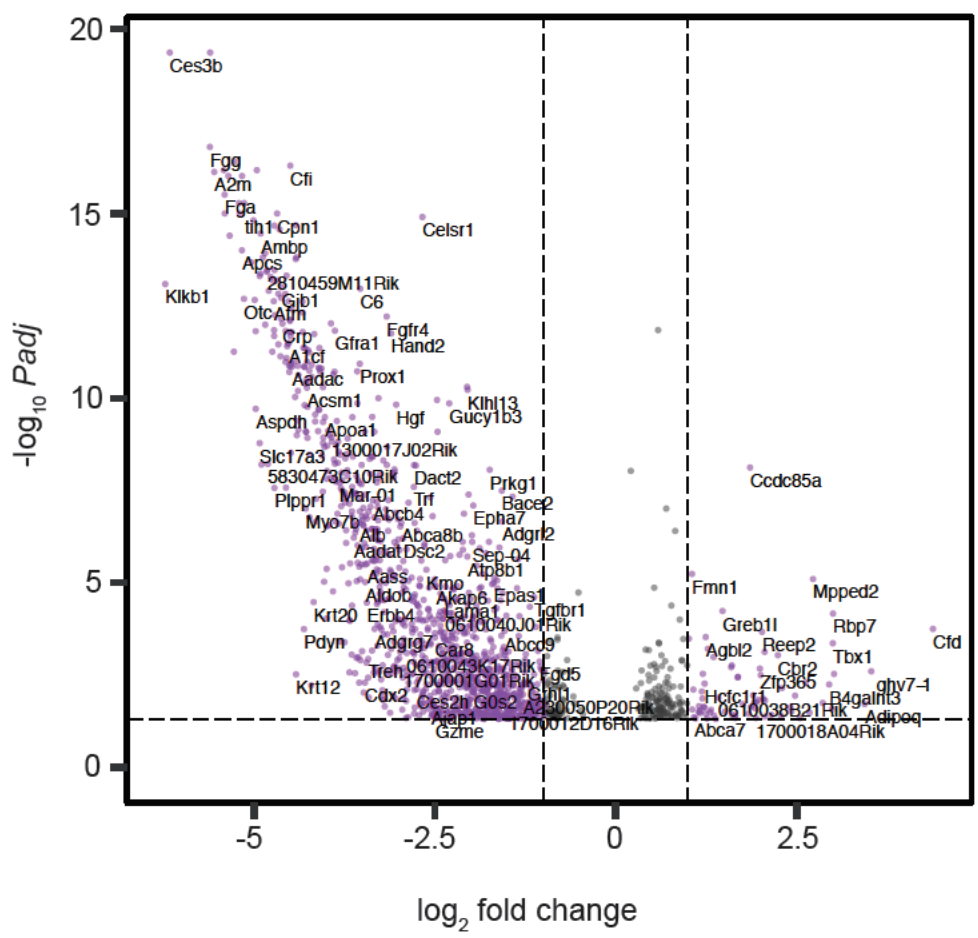


Figure 5.7.4 Principal component analysis of all samples and differential gene expression of Nf2-A compared to the Nf2-negative KRAS^{G12D} screen tumours. A, PCA based on differential gene expression identifies a cluster of samples harbouring Nf2 mutations (Nf2-A, purple ellipse), this group segregates from other samples harbouring Nf2 mutations (Nf2-B). Nf2-A also share Brca2 mutations, and the pair clustering most closely harbour Plk2 mutations. B, differential gene expression between Nf2-A tumours and screen tumours without Nf2 mutations (blue points in PCA)

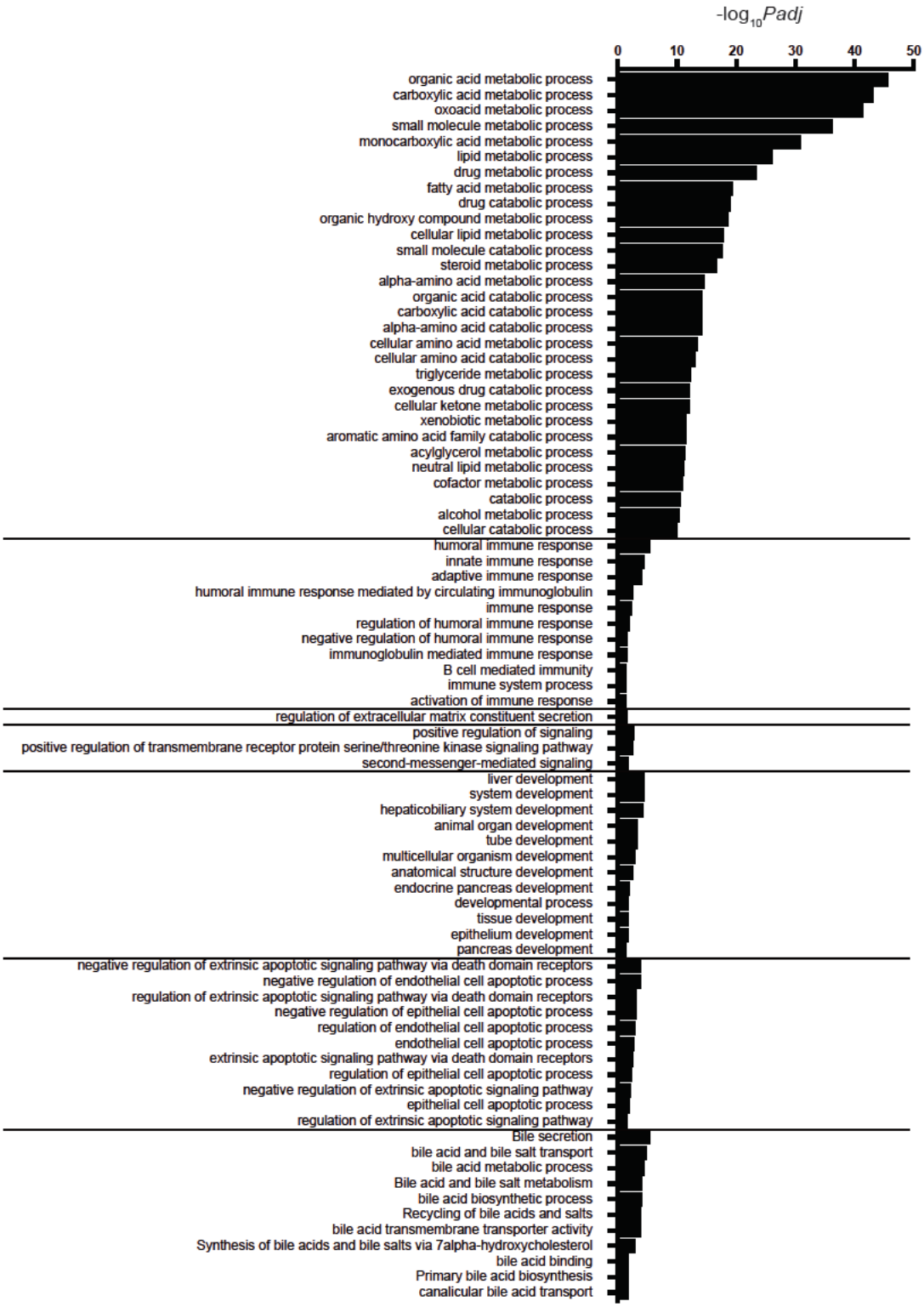

-log10Padj
0
10
20
30
40
50
organic acid metabolic process
carboxylic acid metabolic process
oxoacid metabolic process
small molecule metabolic process
monocarboxylic acid metabolic process
lipid metabolic process
drug metabolic process
fatty acid metabolic process
drug catabolic process
organic hydroxy compound metabolic process
cellular lipid metabolic process
small molecule catabolic process
steroid metabolic process
alpha-amino acid metabolic process
organic acid catabolic process
carboxylic acid catabolic process
alpha-amino acid catabolic process
cellular amino acid metabolic process
cellular amino acid catabolic process
triglyceride metabolic process
exogenous drug catabolic process
cellular ketone metabolic process
xenobiotic metabolic process
aromatic amino acid family catabolic process
acylglycerol metabolic process
neutral lipid metabolic process
cofactor metabolic process
catabolic process
alcohol metabolic process
cellular catabolic process
humoral immune response
innate immune response
adaptive immune response
humoral immune response mediated by circulating immunoglobulin
immune response
regulation of humoral immune response
negative regulation of humoral immune response
immunoglobulin mediated immune response
B cell mediated immunity
immune system process
activation of immune response
regulation of extracellular matrix constituent secretion
positive regulation of signaling
positive regulation of transmembrane receptor protein serine/threonine kinase signaling pathway
second-messenger-mediated signaling
liver development
system development
hepaticobiliary system development
animal organ development
tube development
multicellular organism development
anatomical structure development
endocrine pancreas development
developmental process
tissue development
epithelium development
pancreas development
negative regulation of extrinsic apoptotic signaling pathway via death domain receptors
negative regulation of endothelial cell apoptotic process
regulation of extrinsic apoptotic signaling pathway via death domain receptors
negative regulation of epithelial cell apoptotic process
regulation of endothelial cell apoptotic process
endothelial cell apoptotic process
extrinsic apoptotic signaling pathway via death domain receptors
regulation of epithelial cell apoptotic process
negative regulation of extrinsic apoptotic signaling pathway
epithelial cell apoptotic process
regulation of extrinsic apoptotic signaling pathway
Bile secretion
bile acid and bile salt transport
bile acid metabolic process
Bile acid and bile salt metabolism
bile acid biosynthetic process
Recycling of bile acids and salts
bile acid transmembrane transporter activity
Synthesis of bile acids and bile salts via 7alpha-hydroxycholesterol
bile acid binding
Primary bile acid biosynthesis
canalicular bile acid transport

Figure 5.7.5 Curated biological process enrichment terms in differentially downregulated genes in the Nf2-A group compared to Nf2-negative screen tumours. Marked enrichment for metabolic processes (top) characterises this set of downregulated genes. Also notable is downregulation of physiological processes involved in bile acid homeostasis (bottom). Downregulation of genes involved in immune response, regulation of ECM, liver/biliary development, and apoptosis are seen.

and extensive stroma in these tumours. Interestingly, terms linked to both liver, biliary, and tube development as well as bile acid biosynthesis and secretion strongly suggest a differential loss of biliary cell identity in this group of Nf2 tumours compared to the rest of the screen. This is further supported by enrichment of transcription factor binding motifs for HNF1b, and HNF4a (responsible for maintaining cholangiocyte and hepatocyte lineage identity, respectively) in the set of differentially downregulated genes. Therefore, it might be expected that Nf2-mutant tumours are poorly differentiated and more aggressive.

5.8 *In vivo* validation confirms Plxnb2 and Nf2 as novel drivers cooperating with oncogenic RAS^{G12} isoforms.

Ascribing any definitive driver signal to a particular mutation is hampered by the noise of genetic heterogeneity intrinsic to this screening process. Therefore, following a literature search of all mutations and in light of biological process enrichment and transcriptome sequencing, 5 proposed drivers were chosen for *in vivo* validation: Nf2, Plxnb2, Plk2, Rnf31, and Fh1.

As discussed in the Introduction, Nf2 is an upstream regulator of the Hippo pathway which itself has known roles in iCCA. It has also been shown to have decreased expression histologically in human samples[90]. Plxnb2 and Plk2 shared overlapping biological process annotations and clustered closely to Trp53 (fig. 5.7.3; A). Plk2 mutation is also likely to be the causative mutation explaining the close clustering of samples in the top right of the PCA plot (fig 5.7.4; A). Plk2 is a mitotic kinase and direct transcriptional target of p53, playing a role in cell-cycle progression and centrosome maturation. There is conflicting evidence for its role in cancer. In some studies, including those regarding iCCA, Plk2 expression is anti-apoptotic and pro-proliferative[393,394]. Other studies have shown that Plk2 loss can confer resistance to platinum and taxane-based chemotherapies[395], and it has been implicated as tumour suppressors in gastric[396 398] and hepatocellular carcinomas[399]. Due to its recurrent mutation here, it was appropriate to investigate its potential involvement further. Plxnb2 is a transmembrane semaphorin receptor with roles in neural and cardiac development[400], and immune homeostasis[40] . Interestingly, it

contains an intracellular domain with RasGAP homology and protein structure resembling that of NF1 and RASA1[402] (also mutated in the screen, fig 5.7.2; A, top bar chart). Plxnb2 loss due to epigenetic silencing has been implicated as marker for poor prognosis in lung adenocarcinoma[403], where oncogenic KRAS also plays a crucial role[404]. Rnf31 is the catalytic component of the linear ubiquitin chain assembly complex (LUBAC)[405] which is indispensable for the ubiquitination of NFκB signalling its loss inhibiting the induction of TNF-α induced apoptosis[406]. Indeed, it has been implicated as a tumour suppressor in HCC[407]. Lastly, Fh1 encodes fumarate hydratase, a Krebs cycle enzyme which catalyses the conversion of fumarate to malate. Fh1 is a well-defined tumour suppressor in other cancer types, loss of which leads to metabolome and epigenome remodelling[408]. In a multistep progression, Fh1 loss initially results in senescence but also mediates its escape through indirect hypermethylation of p16[408]. Given the central role of IDH1/2 neomorphic mutations in metabolome/epigenome remodelling in iCCA, it was plausible that Fh1 loss may also act as a tumour suppressor through related mechanisms.

To validate the driver status of these genes *in vivo*, mice were injected via HDTVI such that for each proposed driver, 3 sgRNAs targeting the gene were injected alongside NRAS$^{G\ 2V}$ alone, or NRAS$^{G\ 2V}$ plus 3 sgRNAs targeting Trp53 (sgTrp53). NRAS$^{G\ 2V}$ plus sgTrp53 guides were used as a positive control. For this initial experiment, 3 mice were used per group.

Mice with NRAS$^{G\ 2V}$/driver guides alone did not form tumours within the 10 weeks of the experiment, and those receiving NRAS$^{G\ 2V}$/sgTrp53 harboured some small

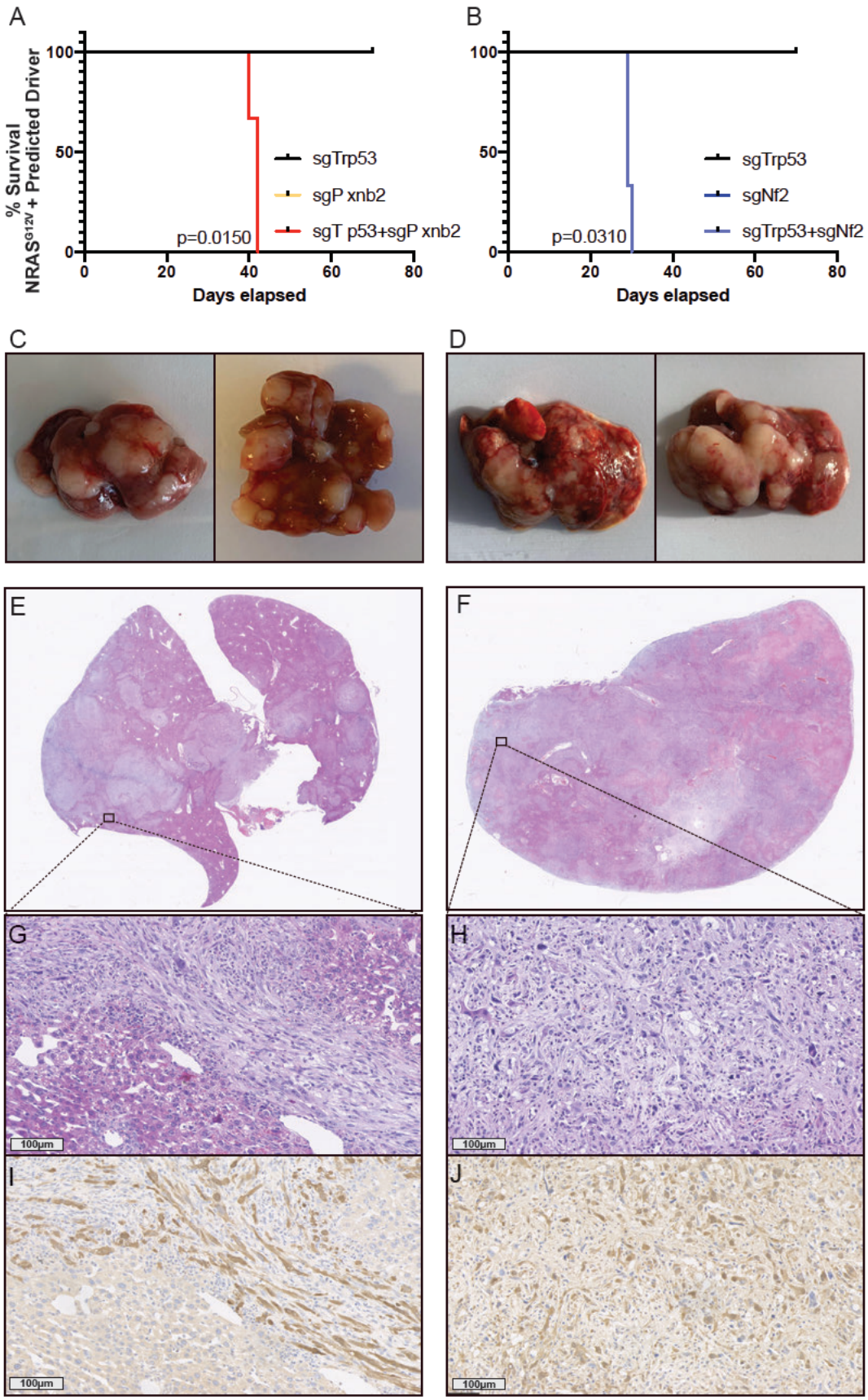

A
% Survival
NRAS^{G12V} + Predicted Driver
100
50
0
0
20
40
60
80
Days elapsed
p=0.0150
sgTrp53
sgP xnb2
sgT p53+sgP xnb2
B
100
50
0
0
20
40
60
80
Days elapsed
p=0.0310
sgTrp53
sgNf2
sgTrp53+sgNf2
C
D
E
F
G
H
I
J
100µm
100µm
100µm
100µm

Figure 5.7.6 Nf2 and Plxnb2 *in vivo* validation results. A-B, Kaplan-Meier survival curves show significantly decreased time to death mice which received sgPlxnb2 or sgNf2 alongside sgTrp53 and $NRAS^{G\ 2V}$, but not with $NRAS^{G\ 2V}$, or alone. C, Macroscopic images of two different livers from $NRAS^{G\ 2V}$/sgTrp53/sgPlxnb2 mice or $NRAS^{G\ 2V}$/sgTrp53/sgNf2 mice (D). E, H&E staining of a $NRAS^{G\ 2V}$/sgTrp53/sgPlxnb2 liver showing extensive tumour spread across the lobe, which is more pronounced in $NRAS^{G\ 2V}$/sgTrp53/sgNf2 mice (F) where almost the entire lobe is consumed by tumour. G-H, H&E, 20x magnification of section enclosed in box in the above overview image. I-J, corresponding GFP staining in the above section, indicative of $NRAS^{G\ 2V}$ expression. GFP can be seen staining spindle-like cells which have lost epithelial morphology.

solid lesions at 10 weeks. One mouse which receiving NRAS$^{G\ 2V}$/sgTrp53/sgFh1 did harbour one tumour at 10 weeks. However, mice that had received NRAS$^{G\ 2V}$/sgTrp53/sgPlxnb2 or NRAS$^{G\ 2V}$/sgTrp53/sgNf2 showed substantial tumour burden with a significant decrease in survival time (P=0.015, NRAS$^{G\ 2V}$/sgTrp53/sgPlxnb2; P=0.031, NRAS$^{G\ 2V}$/sgTrp53/sgNf2) requiring early culling. Macroscopically and histologically, the NRAS$^{G\ 2V}$/sgTrp53/sgPlxnb2 tumours had better defined borders than the NRAS$^{G\ 2V}$/sgTrp53/sgNf2 group, which were very invasive: tumours having taken over almost the entirety of the lobe, with no clear boundaries where one tumour stops and another begins (fig 5.7.6; C-F) Both showed marked loss of cuboidal epithelial morphology and exhibited spindle cell morphology. GFP staining for NRAS$^{G\ 2V}$-expressing tumour cells showed invasive fronts with spindle cells intercalating with morphologically normal surrounding hepatocytes in NRAS$^{G\ 2V}$/sgTrp53/sgNf2 tumours.

Together these results establish that Nf2 and Plxnb2 are novel drivers of intrahepatic cholangiocarcinoma through cooperation with oncogenic RAS$^{G\ 2}$ isoforms and p53 loss.

5.9 Discussion

The identification and validation of driver genes requires a flexible platform that is genetically tractable. The above work utilised the popular hydrodynamic tail vail injection method for delivery of a screening library to the murine liver. This approach initially showed that replacement of myrAKT with the screening library in the commonly used NICD/myrAKT iCCA model increased the onset of lesions by 10 weeks compared to the study initially describing the model[99].

Screening alongside oncogenic $RAS^{G\ 2}$ isoforms uncovered a common set of 27 genes, a strong biological signal shared between screening experiments. These genes showed biological process enrichment for developmental processes and those involved in cell motility. Although not significant at this sample size, it is interesting to note that common iCCA drivers segregated with $NRAS^{G\ 2V}$ and $KRAS^{G\ 2D}$ mutations: ARID1A was lost in three samples exclusively in the $NRAS^{G\ 2V}$ screen, whilst BAP1 and PBRM1 were lost in one sample each in the $KRAS^{G\ 2D}$ screen. Another puzzling observation is the loss of oncogenes in this screen, which preferentially identifies tumour suppressors through loss of function. In the $KRAS^{G\ 2D}$ screen, several kinases which are known oncogenes were mutated, sometimes recurrently: Kras itself, as well as Braf, Erbb2 and 3, and Pik3ca were all mutated preferentially in the $KRAS^{G\ 2D}$ screen. Conversely, Idh1 was lost in the $NRAS^{G\ 2V}$ screen. Whether or not this is a genuine signal or these mutations are passengers is not clear, and warrants further investigation.

Validation of 5 proposed drivers identified Nf2 and Plxnb2 as drivers of iCCA. These mutations occur rarely in the human population: in 1.8% and 1.5% of patients, respectively (n=409, cBioPortal). Here they are shown to significantly accelerate tumourigenesis and mortality through cooperation with both NRAS and KRAS when p53 is lost in the hepatocyte compartment. Both RAS$^{G\ 2}$ oncogenes have been shown to induce senescence in the liver, and this is dependent on p53[243,244]. As such Loss of p53 is required for escape of senescence and is also likely responsible for hepatocyte-cholangiocyte transdifferentiation due to its role in gatekeeping lineage plasticity in the liver.

The contribution of Plxnb2 and Nf2 to the tumour phenotype is the subject of ongoing work. Plxnb2 is a semaphorin receptor, and contains 3 notable binding domains: a GTPase activating protein (GAP) domain with homology to RasGAPS, a Rho-GTPase binding domain (RBD), and a PSD-95/Discs-large/ZO-1 (PDZ) binding domain[409]. Whether or not the Plxnb family have intrinsic GAP activity has been the subject of much research but still remains largely unanswered. Plxnb1 has been shown to exhibit GAP activity on R-Ras and Rap1, but whether or not this extends to Plxnb2 is not yet known[4 0]. Through its PZD binding domains it interacts with PDZ-RhoGEF, and leukaemia associated RhoGEF (LARG) to transduce signals to RhoA and thus mediate cytoskeletal rearrangements in response to semaphorin signalling. Through its RBD it interacts with the atypical Rho-GTPase family members Rnd1/2/3 which remodel the actin cytoskeletin[4 0]. Rnd3 plays a crucial role in inhibiting tumour cell invasion, its activation causes the loss of actin tension and results in cell rounding. Work in melanoma[4], glioma[4 2], and various cancer

cell lines[43,44] has demonstrated that Rnd3 activity suppresses invasive phenotypes/ameboid movement, and that its loss enhances invasiveness. A study by McColl *et al.*[45] showed that Plxnb2 is the effector of Rnd3 through direct binding. They demonstrated that overexpression of Plxnb2 and Rnd3 together, but not alone enhanced inhibition of invasion and cell rounding phenotypes. They finally demonstrated that mutation of the Plxnb2 binding site for Rnd3 abrogated the rounding phenotype and enhanced invasiveness and migration. In light of this, it is possible that loss of Plxnb2 in $NRAS^{G12V}$/sgTrp53/sgPlxnb2 tumours enhances aggressiveness by potentiating invasion and migration, this is supported by the spindle cell morphology seen in figure 5.7.6.

Aside from its previously discussed role in iCCA through dysregulation of the Hippo pathway (see section 1.3.2), Nf2 may also influence phenotype through heterotypic interactions with the Rho-family GTPase RAC1. RAC1 is negatively regulated by Nf2, and Nf2 loss increases activity of RAC1[46]. RAC1 has a plethora of roles in cancer, notably including invasion and migration, increased proliferation, and chemotherapeutic resistance[47]. Overactivation of RAC1 has been shown to promote epithelial-mesenchymal transition in colorectal cancer[48] and is associated with aggressiveness and poor prognosis in both breast[49] and gastric cancer[420]. Myant *et al.*[42] demonstrated a critical role of RAC1 in Wnt-mediated intestinal transformation following *APC* loss. A subsequent study showed that fibronectin-induced cell migration in breast cancer was dependent on Wnt signalling, and that this was mediated through direct interactions between β-catenin and RAC1[422]. This points to RAC1 as an important mediator of proliferation and invasion and this may

be potentiated thought loss of Nf2 and overactivation of canonical Wnt signalling as is seen in iCCA. Therefore, in addition to dedifferentiation due to loss of Hippo signalling, Nf2 loss may synergise with canonical Wnt signalling through release of RAC1 inhibition to drive poorly differentiated, invasive tumours.

Nf2 has previously been explored as a tumour suppressor in the liver. One study showed that its loss in embryonic liver resulted in oval cell expansion and eventual HCC and iCCA. A second study inspired by this explored NF2 mutation and expression in human HCC and iCCA histological sections, finding a low frequency of mutations and low expression in a subset of samples. To my knowledge, the present study is the first example of Nf2 being unbiasedly identified as a driver of iCCA directly from human data. Nf2 was identified through both computational screening of patient exome sequencing and then through *in vivo* screening as part of a library, and finally validated in the adult murine liver. This provides strong evidence in support of what has previously been an incompletely resolved involvement for NF2 loss in human iCCA.

Chapter 6 Discussion and Concluding Remarks

The present study set out to identify novel drivers in iCCA, and in particular, those which have not been detected previously due to their low frequency in the patient population. Leveraging the combined wealth of readily available exome sequencing data served to increase sample size and therefore statistical power to detect uncommon drivers. Indeed, Nf2 was predicted to have driver status by the driver prediction tool MutSigCV which assigns driver status based on mutation frequency in a cohort. Crucially, it was only identified by MutsigCV when the entire pooled cohort was analysed together, not when any one exome-seq study was used alone. In addition to increasing sample size, the use of multiple tools with different prediction models increased the set of candidate drivers. Plxnb2 was not identified by MutSigCV, but by OncoDriveFM, a tool that assigns driver status based on the ratio of benign to damaging mutations; whilst accuracy is increased with sample size, mutation occurrence above two instances is not required.

Screening the set of candidate drivers *in vivo* alongside $KRAS^{G\,12D}$ and $NRAS^{G\,12V}$ oncogenes identified a common set of overlapping drivers. The validation of Nf2 and Plxnb2 as cooperating partners with both oncogenic RAS isoforms and p53 loss demonstrated the validity of this approach. Further exploration of the contexts in which and mechanisms by which these mutations operate will yield deeper insight into the biology of iCCA. Indeed, whether or not oncogenic RAS is an essential requirement for Nf2/Plxnb2 driver function was not explored here, and a follow-

up experiment with Nf2/Plxnb2 loss alongside p53 loss alone or with other oncogenes will shed light on this. Moving forward, it will be interesting to explore the other mutations that were found to be shared between RAS$^{G\ 2}$ screens as well as those found independently with NRAS$^{G\ 2V}$ and KRAS$^{G\ 2D}$. The latter sets are of particular interest and it will be exciting to see if any of them function as drivers only when co-mutated with the respective RAS$^{G\ 2}$ isoform they were discovered with.

In vivo screening explored the oncogenic capacity of prospective human iCCA drivers using a hepatocyte cell of origin. As iCCA can also arise from cholangiocytes a suitable next step will be to develop a parallel screening system which targets cholangiocytes. A possible system could use an adenoviral sgRNA delivery system[423,424] coupled with biliary restricted Cas9 driven from, for example, a CK19 promoter. Further biliary specificity could be achieved through injection of adenovirus into the biliary tree: a recent study described a method for delivery into the murine bile duct through injection into the gall bladder with simultaneous clamping of the common bile duct[425].

A possible disadvantage of *in vivo* screening is the potential for selection of Cas9-low cells biasing the output of the screen. An alternative screening system would be human or mouse biliary organoids. Indeed, this was attempted in mouse biliary organoids using the same libraries as here but packaged in lentivirus for delivery (data not shown). This system offers certain advantages; notably the ability to target cholangiocytes specifically, and also the ability to exert a degree of control over the dose of viral particles (hence the number of sgRNAs) delivered to each cell.

The primary problem is the lack of a clear selection pressure for tumour suppressors or a clear phenotypic output. Furthermore, the characteristic involvement of the acellular stroma in iCCA further excluded the use of organoids as the ECM substrate used to grow them might further bias the results by providing an inauthentic ECM. Another potential solution would be to perform the screen in immunocompromised mice. This would avoid Cas9-mediated immune biasing of the output, but it may also bias the output away from genes involved in immune evasion in the context of human cancer. The cellular stroma is also rich in infiltrating immune cells; indeed, this was shown to be recapitulated in the screen tumours by the presence of regulatory T-cells, neutrophils, and tumour-associated macrophages. Ultimately, the most desirable system is one which recapitulates the complex environment of human tumourigenesis. This necessitates a model with a tissue microenvironment (stroma, blood supply, tissue architecture, mechanical forces), and a systemic environment (circadian cycle, hormonal fluctuations, microbiome) mirroring the multi-scale physiological interplay inherent to human cancer and this includes an immune system. The effect of anti-Cas9 mediated immune selection has not been quantified in the literature, limiting the ability to make an accurate decision about the impact of this on experimental design. With this in mind it was most appropriate to move forward with the screen in immunocompetent adult mice. RAS$^{G\ 2}$-library tumours showed significantly lower Cas9 expression than Cas9-expressing organoids, or NICD/SB-CRISPR-Lib tumours. However, the recurrence of a set of drivers found across tumours, and a subset shared between RAS$^{G\ 2}$ oncogenes suggests that the selective advantage of these mutations outweighs the selective advantage of low Cas9 expression: that following

selection for Cas9-low cells, a genuine biological signal for adaptive driver acquisition persists.

Whilst the primary aim of this study was to identify drivers, the screening model described here may have other useful applications. From a histopathological perspective, compared to the widely used NICD/myrAKT model, $RAS^{G\ 2}$-library tumours more broadly recapitulate morphological and stromal features of advanced human iCCA. NICD/myrAKT tumours are small and well differentiated cystadenocarcinomas that lack the extensive desmoplastic stroma found in classical human iCCA. They are powerful as a model due to their consistency, ease of generation and well-defined molecular/genetic aetiology. The tumours generated here show morphology consistent with ductular, high-grade, and sarcomatoid subtypes, as well as extensive stroma and a broader repertoire of immune infiltrating cell types. As such they represent a complimentary model for the study of the cellular and acellular stroma. Additionally, linking CRISPR/Cas9-induced mutations in this study to different phenotypes will likely yield multiple models with known genetic aetiologies linked to desired stromal or morphological features. This would facilitate the study of a broader spectrum of iCCA phenotypes than is currently possible using existing models.

The screening library may confer great value as a preclinical model for targeted therapeutic response by providing a source of genetic heterogeneity. The inclusion of a library containing human predicted drivers, in combination with a therapeutic target mutation (oncogenic RAS, neomorphic IDH, constitutively active FGFR2 fusions), may constitute a powerful model for dissecting the genetic factors which

contribute to differential treatment response in clinical trials. This represents a logical next step in the development of this system. For the purposes of unbiased driver screening, hypermutated samples were removed at the computational driver prediction stage to decrease false positives. However, when they were included, a number of likely candidate drivers were predicted alongside the proliferation of high-confidence false positives. Although not appropriate in this study due to the risk of biasing the input set for experimental driver screening, manual curation of the driver prediction results with and without hypermutated samples could yield a set of target genes which more extensively model human genetic heterogeneity. This would be advantageous in an experimental model to identify mutational (contra)indications for targeted therapy: a beneficial prospect for a tumour type which is consistently defined by heterogeneity at all levels.

The data sources used as the raw input for this study arise from publications which have contributed immeasurably to the field through the elucidation of common drivers of iCCA. This study aimed to extend this work by using the same data to explore the long tail of infrequently mutated genes in iCCA. The primary result of this study demonstrates that within the noise of genetic heterogeneity exists disease-relevant genes, the mutation of which interacts with more common drivers to influence tumour phenotype and disease progression. Building upon the methodology and results of this study will hopefully shed further light on the genetic aetiology of iCCA and other cancer types more generally.

This work newly identifies PLXNB2 and confirms NF2 as iCCA tumour suppressors which cooperate with common drivers to accelerate disease progression.

Appendix 1: sgRNA spacer sequences

Library-A		
Gene	sgRNA ID	sgRNA Spacer Sequence
Abl1	Abl1_1	CCCTTCATACCGCAGCGAGA
Abl1	Abl1_2	AAGGGAGGGTGTACCACTAC
Abl1	Abl1_3	CGAGAGCCGCTTCAACACTC
Ablim1	Ablim1_1	GTCATAGATCGCCTTGACCT
Ablim1	Ablim1_2	GAAAGTATCTATTCTAGACC
Ablim1	Ablim1_3	CTTACAGTCGGGATGCCACA
Acvr1b	Acvr1b_1	GTGTCTACCATAACCGCCAG
Acvr1b	Acvr1b_2	GTCTACGACCTCTCCACGTC
Acvr1b	Acvr1b_3	AGGGCCGGTTCGGGGAAGTA
Akap9	Akap9_1	GATGCCGTCTCGCTGTTTAA
Akap9	Akap9_2	GCCGCCATTAAACAGCGAGA
Akap9	Akap9_3	GCGCACCTGCTTCAGCCCTT
Apbb1	Apbb1_1	CGAGACGGATTCCGATCTAC
Apbb1	Apbb1_2	GGAATCCGTCTCGAAAGCGT
Apbb1	Apbb1_3	GAAGGACAATGTCGCCTCCT
Aqp7	Aqp7_1	ACTCCCCAGGCATTCGTGAC
Aqp7	Aqp7_2	GGCATCCTTGTTACCGTCCT
Aqp7	Aqp7_3	CATTCGGCCTAGTGCACAAT
Arhgap8	Arhgap8_1	GGCCCTCACCGCACTACTTC
Arhgap8	Arhgap8_2	TCCTCCAGGGGATGACCGCC
Arhgap8	Arhgap8_3	TATAGTCGTTCTCCACATGC
Arid1a	Arid1a_1	GCCCTGCTGGCCATACGCAC
Arid1a	Arid1a_2	GGTCCCTGTTGTTGCGAGTA
Arid1a	Arid1a_3	TAACCCATACTCGCAACAAC
Arid2	Arid2_1	TACCAGCAACACAGCGTGTC
Arid2	Arid2_2	ATAGCGAAGTCCACTTCATT
Arid2	Arid2_3	CACTTTACTGCTCGCTAATG
Atf7ip	Atf7ip_1	CCGTACGTACTTCTCACAGG
Atf7ip	Atf7ip_2	TCCAAAACGTTTGGTTAACC
Atf7ip	Atf7ip_3	GGCAGTGCTCACCGAGCTGC
Atp2b1	Atp2b1_1	TTACGTGCATTATCCCCTTC
Atp2b1	Atp2b1_2	AATTGCAGCCATAGTATCAT
Atp2b1	Atp2b1_3	CCTGCTCAATTCGACTCTGC
Atp7a	Atp7a_1	AACACGGTATTGGTTAAGAC
Atp7a	Atp7a_2	CAAGGTGTTCAGCGCATTAA

Atp7a	Atp7a_3	GCAGCCGAAGTACCTCAAAT
Bap1	Bap1_1	TCAAATGGATCGAAGAGCGC
Bap1	Bap1_2	CCGCCGCAAGGTTTCTACGT
Bap1	Bap1_3	GCCCACGCTGAGCCGAATGA
Bcor	Bcor_1	GATGAGGCCGAATCCAACGA
Bcor	Bcor_2	CTTACGGGCTGCATACGAGC
Bcor	Bcor_3	CGACGCTTCAAAAGCCAGGC
Braf	Braf_1	TCATAGGTACCCGCAAGATG
Braf	Braf_2	GAGGCCCTATTGGACAAATT
Braf	Braf_3	GTTGTTCCGTGATGCATCTG
Brca2	Brca2_1	TAGGACCGATAAGCCTCAAT
Brca2	Brca2_2	AAAGCTCCTCAAACCAATTG
Brca2	Brca2_3	GGACTAGCAACATCTACCAC
Cacna1c	Cacna1c_1	CAACCGCCACACGGCCGCCC
Cacna1c	Cacna1c_2	CTCCTCGCAGGTTCCAACTA
Cacna1c	Cacna1c_3	GTTGGAGTCGTCTTCCGGAA
Camkk1	Camkk1_1	ACAGGCCGTCCTCCCCCCAG
Camkk1	Camkk1_2	TCTCCTGATACACACGCTCC
Camkk1	Camkk1_3	TCTGTCTTCACTTTCGTTGT
Cc2d2a	Cc2d2a_1	ACTCACTCCGGGGCCCCCGC
Cc2d2a	Cc2d2a_2	AGGCAGAAAGCGCATTGTTG
Cc2d2a	Cc2d2a_3	CAACAATGCGCTTTCTGCCT
Ccdc66	Ccdc66_1	TGAATACTGCTTATCCGTAC
Ccdc66	Ccdc66_2	CACTACCGGTACTTGTCCAG
Ccdc66	Ccdc66_3	TGCTCATGTTCACTCGTGAC
Cdk18	Cdk18_1	GCCCTGCACTGCTATTCGAG
Cdk18	Cdk18_2	GAACCAGCGCCGATTCTCCA
Cdk18	Cdk18_3	TCCCAGTCAGATATCGGCTT
Cdkn2a	Cdkn2a_1	GTGCGATATTTGCGTTCCGC
Cdkn2a	Cdkn2a_2	CCCAACGCCCCGAACTCTTT
Cdkn2a	Cdkn2a_3	GGGGTACGACCGAAAGAGTT
Chd4	Chd4_1	GCTTCGGATAAATCCTCGTC
Chd4	Chd4_2	TCGGACCCTCACCAACTACA
Chd4	Chd4_3	TCCTGACTTACCTAACAAAT
Cped1	Cped1_1	TAGTCGACACATGTGCTCCC
Cped1	Cped1_2	TGGTCTAGGTCCGGAGCTGC
Cped1	Cped1_3	GTCGACTAAGCCTACATCAA
Cpt1a	Cpt1a_1	GACTATGCGCTACTCGCTGA
Cpt1a	Cpt1a_2	CACTCACGATGTTCTTCGTC
Cpt1a	Cpt1a_3	GGGTTACTCACGTGGTGTCT
Dbndd2	Dbndd2_1	CATAGGTAGCATCTCGTCTA

Dbndd2	Dbndd2_2	GTTTATTGATCTTGCGGATC
Dbndd2	Dbndd2_3	GATGTGGTCCTGTCAGACGT
Dchs1	Dchs1_1	TAGGTAGGGCTATCAAATCG
Dchs1	Dchs1_2	TGTGTGTACCACCCGGACCC
Dchs1	Dchs1_3	TAGAGCGCGCAGTGTATAGA
Ece2	Ece2_1	GAACGTCGCGCTGCACGAGT
Ece2	Ece2_2	CTCGTGCAGCGCGACGTTCA
Ece2	Ece2_3	GACCATGGATGTTCGAGCAC
Ehhadh	Ehhadh_1	AGCTGCGTTCCTCTTGCACC
Ehhadh	Ehhadh_2	CTATCGGATTGCCAATGCAA
Ehhadh	Ehhadh_3	GCCATCCAAGGCGTGGCTCT
Elf3	Elf3_1	CGACTTCTCCCGCTGCGACA
Elf3	Elf3_2	CTAGCGAGCGGCCCCAGTTC
Elf3	Elf3_3	CAAAGGGGCCCGAGTCGCCT
Epha2	Epha2_1	CAACGTGGTATCCGGCGACC
Epha2	Epha2_2	TTCGCTGTCGAAGCACGCAA
Epha2	Epha2_3	CTGCTGACCGTGATCTCGTC
Erbb2	Erbb2_1	CTTACAGGCCCGGGAACGAT
Erbb2	Erbb2_2	GCTAGACAACCGAGACCCTT
Erbb2	Erbb2_3	ACATGGACACCAATCGTTCC
Erbb3	Erbb3_1	CCCCTTGCAGACTTCGTGAC
Erbb3	Erbb3_2	CGGGGAACCCAGGTCTACGA
Erbb3	Erbb3_3	GCCCTTACCTAACCTCCGAG
Esrra	Esrra_1	TGTACTTCTGCCGTCCGCCG
Esrra	Esrra_2	GTACGTCCTGCTGAAAGCTC
Esrra	Esrra_3	TGCGACACCAGAGCGTTCAC
Etv6	Etv6_1	GCATGGCGTGCTCTTCCGGT
Etv6	Etv6_2	GTGAACATGAAGCGGAGTCG
Etv6	Etv6_3	CTGACAGGGGGTACGTTTCC
Fam136a	Fam136a_1	GCGCCGCTACCTGCATCTTC
Fam136a	Fam136a_2	CGAGCGCTGCCATGCGCCTC
Fam136a	Fam136a_3	CATGGCAGCGCTCGATGCAT
Fam174b	Fam174b_1	CCACGATCACCGTCGCCTTC
Fam174b	Fam174b_2	TCGCAGAGCCGAGTCCGCTT
Fam174b	Fam174b_3	GCCATTTCCACACGTTCCGC
Fh1	Fh1_1	AATTGGGCGAACTCACACGC
Fh1	Fh1_2	CGTGTAGAGTTCGACACCTT
Fh1	Fh1_3	AAAATCCAAAGAGTTTGCGC
Fn1	Fn1_1	CCAGGTCTCCCCCACGACGT
Fn1	Fn1_2	GACCTACCTAGGCAACGCCC
Fn1	Fn1_3	CCTACAAGATTGGCGACAAG

Fzd2	Fzd2_1	ATCGACGGCGACCTGCTGAG
Fzd2	Fzd2_2	GGCGCGGTGGTGAGTAGCGC
Fzd2	Fzd2_3	TGCGAGCATTTCCCGCGTCA
Gpr162	Gpr162_1	AGTCCAACGGCTCGGCTATC
Gpr162	Gpr162_2	TCTCCGGATAGCCGAGCCGT
Gpr162	Gpr162_3	GTGTGTGGCGATCATGTCCG
Gstp1	Gstp1_1	GTCAGCCAGCAGCATTCGCA
Gstp1	Gstp1_2	ATTGTCTACTTCCCAGTTCG
Gstp1	Gstp1_3	TCACATAGTTGGTGTAGATG
Heatr3	Heatr3_1	CACTACCAGACTTAGCCCGT
Heatr3	Heatr3_2	GCCGGCGCACAGCATCCCGA
Heatr3	Heatr3_3	ATTGACCTGGCTGTCTCCGT
Helz	Helz_1	GCAACGAGTAATGTCCTGTC
Helz	Helz_2	CTTCTACGTGACATCCAGAC
Helz	Helz_3	GCTGATGAAGATTGTAGGCA
Hrct1	Hrct1_1	GAAACCGAGTCCGACAAGCC
Hrct1	Hrct1_2	GTGGTGGTGAACGCCCACAC
Hrct1	Hrct1_3	CGCCGGCCATGGAAGAGTCG
Idh1	Idh1_1	TGGGCCTGTAAGAATTACGA
Idh1	Idh1_2	GGCCCAAGCTATGAAGTCCG
Idh1	Idh1_3	AGTCTTCAATTGACTTATCC
Idh2	Idh2_1	ATGTTCCGGATCGTTCCGTT
Idh2	Idh2_2	AGAGCCCTAACGGAACGATC
Idh2	Idh2_3	TCGAGCTGGCACGTTCAAGT
Il1r1	Il1r1_1	AGTCCCGGTCCGCTGATATG
Il1r1	Il1r1_2	GCCGTATGTCCTATACGTTC
Il1r1	Il1r1_3	TTGCTTCCCCCGGAACGTAT
Itpr2	Itpr2_1	GGACATCGTGTCCCTGTACG
Itpr2	Itpr2_2	TGACCGAGCCCTCCGCGTAC
Itpr2	Itpr2_3	GTTCCCCTGTTTCGCCTGCT
Itpr3	Itpr3_1	AGCATCCAGCGTCACCCGCA
Itpr3	Itpr3_2	GCTGGATGCTACGGGCAATG
Itpr3	Itpr3_3	GGTCTTACCTCGGAACTTCT
Kras	Kras_1	TGAGTATGACCCTACGATAG
Kras	Kras_2	AGCAGCGTTACCTCTATCGT
Kras	Kras_3	TAGAACAGTAGACACGAAAC
Mme	Mme_1	TGCTCGACTGATTCAGAATA
Mme	Mme_2	GGCCAGTAGCATCAGATAAC
Mme	Mme_3	TCCCAGTTATCTGATGCTAC
Msr1	Msr1_1	GCGTTCCGTGTCTATAAGGT
Msr1	Msr1_2	TGAACGTGCGTCAAATTTCA

Msr1	Msr1_3	TTCCTTGATTTCGTCAGTCC
Mtch2	Mtch2_1	CTTTGACCGAGTTATCAAAG
Mtch2	Mtch2_2	GGCACCTACCATAGCAAAAG
Mtch2	Mtch2_3	AAAATATTTCGTCCTATTGT
Ncor1	Ncor1_1	GGATATGAACAGTTTCACTC
Ncor1	Ncor1_2	AATGTATTAGGCCTCAAGAA
Ncor1	Ncor1_3	TATGCCTTACCTGCTGGTGT
Nf2	Nf2_1	GATCCGCACCGTGAATGTCT
Nf2	Nf2_2	GAAGCTCATGCGAGAAGCGA
Nf2	Nf2_3	CTTGGCGTCATATGCTGTCC
Notch2	Notch2_1	GTCGTCGATATTCCGCTCAC
Notch2	Notch2_2	GCCTCCGTTGACGCAGGGCG
Notch2	Notch2_3	TACGAGTGCACCTGCCAAGT
Nprl3	Nprl3_1	AGCGGCGCTCCTCATGCTGC
Nprl3	Nprl3_2	CTGCCAGTACCTCACTCGAG
Nprl3	Nprl3_3	AAGGTTCTCGGATGTTATTC
Nras	Nras_1	CTAGCATACCTGTCGGGTCT
Nras	Nras_2	GACCTCAGCCAAGACCCGAC
Nras	Nras_3	TTGCAGATATTAACCTCTAC
Pbrm1	Pbrm1_1	CAAACTCATTTCTTGTTCGA
Pbrm1	Pbrm1_2	GTTGTAGCCACAAATCCATC
Pbrm1	Pbrm1_3	ACATCACTTACCACTTTGGA
Phf20l1	Phf20l1_1	CACGATCCAGCTGGGTCGAT
Phf20l1	Phf20l1_2	GATATGATGAGTGGATTTAC
Phf20l1	Phf20l1_3	TGTTGTCCTATCGACCCAGC
Pik3ca	Pik3ca_1	GCGCACTATTTATGACCCAG
Pik3ca	Pik3ca_2	TCACCATGCCGTCATACTCC
Pik3ca	Pik3ca_3	CAGAAGTCCAAGACTTTCGA
Pik3r1	Pik3r1_1	AGATGCGTCTCGTACCAAAA
Pik3r1	Pik3r1_2	GGAGTACACCCGTACTTCCC
Pik3r1	Pik3r1_3	GTGAATTACCTGCTGGTATT
Plch2	Plch2_1	TAGCCCGTCGCCAGCGTACC
Plch2	Plch2_2	CCGTAGTCCCCCAAGCCAAA
Plch2	Plch2_3	GATATTGGTCCCTGGTACGC
Plk2	Plk2_1	CAGAAGTCCGATACTACCTC
Plk2	Plk2_2	AAAACTGCACGACATGCTTA
Plk2	Plk2_3	ACGAACAAGAAATCTTGCAC
Plxnb2	Plxnb2_1	CTCAGATGGCCGGATCCTTA
Plxnb2	Plxnb2_2	TACTCTGCAGAAGTGCCGTC
Plxnb2	Plxnb2_3	GGAGTCACGACACTGAGCGC
Rasa1	Rasa1_1	ACACGCCTTCTATCTTCTAC

Rasa1	Rasa1_2	CTATAGCAGAAGAACGCCTC
Rasa1	Rasa1_3	CGTGACTGTAATAACCTATT
Rb1	Rb1_1	AGAAATCGATACCAGTACCA
Rb1	Rb1_2	TGACATAGCATTATCAACCT
Rb1	Rb1_3	TTGGGAGAAAGTTTCATCCG
Rbmx	Rbmx_1	CCTTAAGCCCCTGTGTCCCG
Rbmx	Rbmx_2	CTCTTGACTTATTCGTTTCT
Rbmx	Rbmx_3	CGAGAAACGAATAAGTCAAG
Rev1	Rev1_1	ACGGCGTCTGCAATATTTAG
Rev1	Rev1_2	TGAAATGAAGTTGCAGTCCG
Rev1	Rev1_3	TTGGTCGTAGGGCGGGTACA
Rexo4	Rexo4_1	TGATACTTACTCAGTGGGCG
Rexo4	Rexo4_2	CCCTAGCTTGACAAAAGCCT
Rexo4	Rexo4_3	GTGTCGATCGTGAACCAGTA
Rnf31	Rnf31_1	CATACAACCGTAGTACATCC
Rnf31	Rnf31_2	AGGGTGGCCGGGATGTACTA
Rnf31	Rnf31_3	TAACCCCGTCTTTCGCAGCA
Rnf43	Rnf43_1	GAGACGCTTACCCCGGCGGG
Rnf43	Rnf43_2	CAGGGGCGAGGAGCTCGTCG
Rnf43	Rnf43_3	TTTCCACAGGCCCGAATGGC
Rps29	Rps29_1	CTCACCAAGAGCGGGAACCC
Rps29	Rps29_2	GAAGGACATAGGCTTCATTA
Rps29	Rps29_3	TGGGTCACCAGCAGCTCTAC
Sdk1	Sdk1_1	AGAACTCACGTGCTACGCTC
Sdk1	Sdk1_2	TCCTGTTCCGCACCACGCAG
Sdk1	Sdk1_3	CATCACTCGCTGTTGTAGCA
Setd2	Setd2_1	GATCTCTTTCGGACCGACAT
Setd2	Setd2_2	GTCGGTCCGAAAGAGATCGA
Setd2	Setd2_3	ACTGCATTCGCTTAATATCC
Sf3b1	Sf3b1_1	GTCCTCCAAAGATTGCCGAT
Sf3b1	Sf3b1_2	GCCTGGATATCATGCCCCCG
Sf3b1	Sf3b1_3	TATATCATTAAGCAACGCCA
Smad4	Smad4_1	GCCAAGTAATCGCGCATCAA
Smad4	Smad4_2	TCCGTTGATGCGCGATTACT
Smad4	Smad4_3	ACAACCCGCTCATAGTGATA
Smu1	Smu1_1	TCCAGATGAATATAGCGCTC
Smu1	Smu1_2	ATTGATAGAGCTTCGTGAAT
Smu1	Smu1_3	CTTGCCTCGAAACAAATCGA
Snrpn	Snrpn_1	GGAACTCCACCTCCACCTGT
Snrpn	Snrpn_2	TGGGGAATAGGTACACCTGC
Snrpn	Snrpn_3	TTCACAGGTCATGACCCCAC

Stk11	Stk11_1	GCGCCCTACGTATATGGTGA
Stk11	Stk11_2	TGTACAGCACGTCCACAAGC
Stk11	Stk11_3	ATTCCAGGCCGTCAATCAGC
Tbc1d8b	Tbc1d8b_1	CTAGGGACTAATCGCTGAAG
Tbc1d8b	Tbc1d8b_2	GTCTGGTGTAGGATGCGAAA
Tbc1d8b	Tbc1d8b_3	CCGTCCTCTCCAGTAACTGC
Tgfbr1	Tgfbr1_1	ATCTATTCAAGTAATCGAAA
Tgfbr1	Tgfbr1_2	CTTCTAGAGAAGAGCGTTCA
Tgfbr1	Tgfbr1_3	AGTGATGGATCCTCTTCATT
Thbs1	Thbs1_1	AAGGGGCCCCGGTCGCCGAC
Thbs1	Thbs1_2	CCCTTCACCAGTCGGCGACC
Thbs1	Thbs1_3	GCTATCCGCACCAACTACAT
Tmtc1	Tmtc1_1	GTAGAATGACATTGACCGCG
Tmtc1	Tmtc1_2	TTATACTCACCGCCTCAGTG
Tmtc1	Tmtc1_3	CCATGTCGCCGAGAGCTATG
Tmtc4	Tmtc4_1	GGCCTTTACCGGTGTACTGT
Tmtc4	Tmtc4_2	AGGATTAACTACTACCTGTC
Tmtc4	Tmtc4_3	CCCCAAGGGGCGTGTCTGAC
Trp53	Trp53_1	AGTGAAGCCCTCCGAGTGTC
Trp53	Trp53_2	AACAGATCGTCCATGCAGTG
Trp53	Trp53_3	TGAGGGCTTACCATCACCAT
Ttc8	Ttc8_1	CGATCTATGCACGCAGATGC
Ttc8	Ttc8_2	CGACCTATCACTAGCTCATC
Ttc8	Ttc8_3	CGGCCAGGTACCTGATCATA
Wdr62	Wdr62_1	ATGACATGGTCCTCAATGTT
Wdr62	Wdr62_2	GTCACTATGTACTTCCCATC
Wdr62	Wdr62_3	ATGTGCTTACCCTGCTAAGT
Zfp317	Zfp317_1	CCATAAATAGGTTATCAGGT
Zfp317	Zfp317_2	TTATAGCAATCTAAGTTCAC
Zfp317	Zfp317_3	GAACTACATCTAAAGTCAAA
Zfp36l2	Zfp36l2_1	CGATATCGACTTCTTGTGCA
Zfp36l2	Zfp36l2_2	CGCGTGCGCCAAACGCTCGC
Zfp36l2	Zfp36l2_3	GCCGCCGCTCGTCCGCGTTG

Library-B		
Gene	sgRNA ID	sgRNA Spacer Sequence
Abl1	Abl1_1	CGCAACAAGCCCACTATCTA
Abl1	Abl1_2	TCATCACCGCCGCTTCCTTC
Abl1	Abl1_3	GTCTCAGCGAAGCAGCTCGA
Ablim1	Ablim1_1	TCTTTCCAGGTTCTACCGTG
Ablim1	Ablim1_2	TATGAGCCTTTCTACACATC
Ablim1	Ablim1_3	TCCAGCATCCCTGGTTCACC
Acvr1b	Acvr1b_1	CACGCCATACTTCCCCGAAC
Acvr1b	Acvr1b_2	CATGGTGCTCCACGCCATCC
Acvr1b	Acvr1b_3	CCGCTACACAGTGACCATTG
Akap9	Akap9_1	GCCGGGAATCCTGATTGCTC
Akap9	Akap9_2	TGATATTGAAGAGTTGACTC
Akap9	Akap9_3	CGAGCTACATTCATCTGCAG
Apbb1	Apbb1_1	CTCCTTCAGCTCACCTGGAC
Apbb1	Apbb1_2	CTTGAACAAACCTTGATCCC
Apbb1	Apbb1_3	CCCAATGGAGTTGGGACTGA
Aqp7	Aqp7_1	GGTGCCATTAACCACTTTGC
Aqp7	Aqp7_2	GAAGTTCCCTGTATATGTGC
Aqp7	Aqp7_3	CAGAGATGCCGCCTGCTACA
Arhgap8	Arhgap8_1	CTTCCGGTCAAACTCCTTGT
Arhgap8	Arhgap8_2	GTACAACGTACAGTGCCTTC
Arhgap8	Arhgap8_3	GCTCAACCACCAGCGCCTGC
Arid1a	Arid1a_1	AGTGTCGCGCAGTCTCGCTC
Arid1a	Arid1a_2	TGAGCGAGACTGCGCGACAC
Arid1a	Arid1a_3	CTGCTGCTGATACGAAGGCT
Arid2	Arid2_1	TACCTGACACGCTGTGTTGC
Arid2	Arid2_2	GCCGTTTAAGAAGATCCCTG
Arid2	Arid2_3	TCCGCCTAAAGTAGTGACTC
Atf7ip	Atf7ip_1	TGCTGTGTTTGATAAGACGT
Atf7ip	Atf7ip_2	ATTCTTCATCCATATATCGC
Atf7ip	Atf7ip_3	AGAGGAGCGAGTTCCAGCCG
Atp2b1	Atp2b1_1	TGGGCTTACCGTACTTCACT
Atp2b1	Atp2b1_2	GGAATCATCTTCACCTTACT
Atp2b1	Atp2b1_3	TGCTATAGGTACTCATGTGA
Atp7a	Atp7a_1	CTTCAGGCGCTCAACATCAA
Atp7a	Atp7a_2	GTTTGCTACACAAGAAGCAC
Atp7a	Atp7a_3	CGCAATTTAAGACGAGAAGA
Bap1	Bap1_1	ATCCTTCATTCGGCTCAGCG
Bap1	Bap1_2	CCTGATCGTAGGTGTCAAAG
Bap1	Bap1_3	TCTTACCGAAATCTTCCACC

Bcor	Bcor_1	CTGTGAACGTTCCCATACAG
Bcor	Bcor_2	GAAGCGTCGCCATCATTCAC
Bcor	Bcor_3	ACAGCTGGATGAACAGCGAG
Braf	Braf_1	GCAAATGATTAAGTTGACAC
Braf	Braf_2	TCTTCCTGCCCAACAAACAG
Braf	Braf_3	CTTGCTGGTGTACTCTTCAT
Brca2	Brca2_1	TGTGTGCACCGCAGAGTAAG
Brca2	Brca2_2	GTACTACAACGAGAAAAGCC
Brca2	Brca2_3	TCACTGAGACAAGATTTCAG
Cacna1c	Cacna1c_1	CCCAACGCTTACCTCCGCAA
Cacna1c	Cacna1c_2	GCCGCAGTGGACGGAGCACG
Cacna1c	Cacna1c_3	TCCGCGTGCTCCGTCCACTG
Camkk1	Camkk1_1	TCTGCAGGGTGCCTATGGTG
Camkk1	Camkk1_2	TACAAGCTGCAGAGTGAGAT
Camkk1	Camkk1_3	TCTCACTCTGCAGCTTGTAC
Cc2d2a	Cc2d2a_1	CAAGCTCAGGAGGTTGGTCT
Cc2d2a	Cc2d2a_2	CTCCCTCATTCGTTCCCTCA
Cc2d2a	Cc2d2a_3	AAGCCGGTTCTAGAAGCTGA
Ccdc66	Ccdc66_1	ATATGGGGACAGTATTAACT
Ccdc66	Ccdc66_2	GAAATGCCTCTTGAGCTCTC
Ccdc66	Ccdc66_3	AAGTTAATACTGTCCCCATA
Cdk18	Cdk18_1	ACCCTCTCGAATAGCAGTGC
Cdk18	Cdk18_2	CGTTGTTGATACTGCATGCC
Cdk18	Cdk18_3	CATGGACATCCGCCTACCCC
Cdkn2a	Cdkn2a_1	GGCTGGATGTGCGCGATGCC
Cdkn2a	Cdkn2a_2	TCGTGCGATCCCGGAGACCC
Cdkn2a	Cdkn2a_3	CCGCACCGGAATCCTGGACC
Chd4	Chd4_1	TTGCGGATCGGCACCTCCAC
Chd4	Chd4_2	TAAGAGCAAGCGCCAAAAAA
Chd4	Chd4_3	GGCTCTGCGCTCAGACAGTG
Cped1	Cped1_1	TCCCGAGGTTTCCATACATG
Cped1	Cped1_2	GCAGAAGGGCCGTCCTCCAC
Cped1	Cped1_3	AAGATTACCTGGAAATCTTC
Cpt1a	Cpt1a_1	GCAGGCGCGGCAGAGACGTC
Cpt1a	Cpt1a_2	CTGTCAAAGATACCGTGAGC
Cpt1a	Cpt1a_3	GGGGTCCACTTTGGTATGCA
Dbndd2	Dbndd2_1	ATTTGGACTGCGCAGGTTGG
Dbndd2	Dbndd2_2	GCCGCAGCTGTTGGCGCTCC
Dbndd2	Dbndd2_3	AGCGCCAACAGCTGCGGCTC
Dchs1	Dchs1_1	AGAATTAGCCCGTCTGCCTA
Dchs1	Dchs1_2	ATTATGCCCACACTGGTGCC

Dchs1	Dchs1_3	AAAGTCGCCCATGGGGTCCC
Ece2	Ece2_1	GTCCAGTGCTCGAACATCCA
Ece2	Ece2_2	GTAAGCCGGTTGCTAGTTCC
Ece2	Ece2_3	ATGCTCAATCCCGTTATGAC
Ehhadh	Ehhadh_1	GCTCTTGACTTAATTACCTC
Ehhadh	Ehhadh_2	AACCGTCTGAGCAAATTTGA
Ehhadh	Ehhadh_3	TATGGTCCAAACTAGCTTTC
Elf3	Elf3_1	ACTCCTGGATGATGGCCGCC
Elf3	Elf3_2	CCCTACTACTGCAGTACCTA
Elf3	Elf3_3	CCTGGACCTCACCGAGAGCA
Epha2	Epha2_1	GTGCCAAAGTAGAACTGCGT
Epha2	Epha2_2	CTCACGCACCCCTATGGCAA
Epha2	Epha2_3	TTTACTCACCCCTTTGCCAT
Erbb2	Erbb2_1	TGACGTTGTCCAAAGGGTCT
Erbb2	Erbb2_2	CATGCTTCGCCACCTCTACC
Erbb2	Erbb2_3	AGGGAGTCGCAACTTCATGT
Erbb3	Erbb3_1	CTCAGCCCGTTTAGAGTCCC
Erbb3	Erbb3_2	GTCCCGTAAGCACAATCTCC
Erbb3	Erbb3_3	TTGACTGGAGGGACATCGTG
Esrra	Esrra_1	CCACTATCTCTCGATCAAAA
Esrra	Esrra_2	TACAATTCAAGGTGTGCGTC
Esrra	Esrra_3	GAGCATCGAGTACAGCTGTC
Etv6	Etv6_1	CAGGTTTGCAGCCGATTTAC
Etv6	Etv6_2	ACACAGTGCCTCGAGCGCTC
Etv6	Etv6_3	CGCTACCGATCTCCTCATTC
Fam136a	Fam136a_1	TTGTTGGGATGAGGTGCATG
Fam136a	Fam136a_2	GGTGCAGCAACTCCGAGTGC
Fam136a	Fam136a_3	TGACAAAGCCAAAGACTCAA
Fam174b	Fam174b_1	TCATCGTCCTCCTCATTCAG
Fam174b	Fam174b_2	CTCCAACAGCAGCGTCGACG
Fam174b	Fam174b_3	GTCTGCTGCTGCGCGTCTTC
Fh1	Fh1_1	ATGCCGAAAGCTTGAATGAC
Fh1	Fh1_2	GTTCCAACCGATAAGTATTA
Fh1	Fh1_3	AGCTGCTGCTGAAGTAAACC
Fn1	Fn1_1	CCTACGTCGTGGGGGAGACC
Fn1	Fn1_2	GACTGTACCTGCATCGGGGC
Fn1	Fn1_3	CCACTTACCTATAGGTTTGC
Fzd2	Fzd2_1	GCCGGACCGGCCCAGTTCCA
Fzd2	Fzd2_2	CGGTTATCGCACGGTGGTGC
Fzd2	Fzd2_3	ATCTGCTCCGCGCCGTGACG
Gpr162	Gpr162_1	CCCGTTGTTGTGCCAGCCGA

Gpr162	Gpr162_2	GGGGAGCGCTACTATGCCCG
Gpr162	Gpr162_3	GCACATGCTCTTCCCACCGC
Gstp1	Gstp1_1	CATCATTCACCATATCCATC
Gstp1	Gstp1_2	TCGAACTGGGAAGTAGACAA
Gstp1	Gstp1_3	GCGAATGCTGCTGGCTGACC
Heatr3	Heatr3_1	AAAGGACATCATGACGCCCC
Heatr3	Heatr3_2	TTGTGTTACAGTACTTACGT
Heatr3	Heatr3_3	TGTAGAGAGCATAGCCAATG
Helz	Helz_1	AGACTCCCCTGACAGTGACT
Helz	Helz_2	CTGTGCTGCATGCAGCTGAA
Helz	Helz_3	GACACTTCTCACAGAAACAG
Hrct1	Hrct1_1	TCACCCCGGTCACATGGTGA
Hrct1	Hrct1_2	AGCGGCGGGCTCCATGAGCA
Hrct1	Hrct1_3	CTTCCAGATGTTAGGCCTTC
Idh1	Idh1_1	TCGTGATGCCACCAATGACC
Idh1	Idh1_2	TGGCACCATCCGAAACATTC
Idh1	Idh1_3	ACTATTCTGAAGAAGTATGA
Idh2	Idh2_1	GGGGATGTACAACACCGACG
Idh2	Idh2_2	TGGACGGTGACGAGATGACC
Idh2	Idh2_3	CAAGGTGGAGAAGCCGGTAG
Il1r1	Il1r1_1	CAGCAAGACCCCCATATCAG
Il1r1	Il1r1_2	ACGAAATTACCTTATACCAC
Il1r1	Il1r1_3	TGAAAATAATGAGTTACCCG
Itpr2	Itpr2_1	TTGGCAAGGTCTCCCGCCTC
Itpr2	Itpr2_2	TGGGCAGAGTATCGATTCAT
Itpr2	Itpr2_3	CTTGTAGCATGCTGCAGAGC
Itpr3	Itpr3_1	AACCCCCCCAAGAAGTTCCG
Itpr3	Itpr3_2	CACGGGGACGTGGTGAAGTA
Itpr3	Itpr3_3	TTTGTCCTGTTTAGTCTGCT
Kras	Kras_1	AATCACACTTATTCCCTACC
Kras	Kras_2	CCTCCCCAGTTCTCATGTAC
Kras	Kras_3	TCTCGACACAGCAGGTCAAG
Mme	Mme_1	ATTTATACACGATCTGTACA
Mme	Mme_2	GACTTCTTCAAGTATGCTTG
Mme	Mme_3	TCATCAGACTGCATAAAATC
Msr1	Msr1_1	GCAGTACTAATACCTGTTGT
Msr1	Msr1_2	GGAGAAGTTGAAGTCCTTCA
Msr1	Msr1_3	TCAGCATCCTCTTGTTCATG
Mtch2	Mtch2_1	TTGACTCCAAGACTGTGTTC
Mtch2	Mtch2_2	TTACCTGTAAGACTTTCCCA
Mtch2	Mtch2_3	CTCTTACCTCTTTGATAACT

Ncor1	Ncor1_1	AGTTAATACAGAGTATGGAC
Ncor1	Ncor1_2	TCCCAGTACCCGACACCAGC
Ncor1	Ncor1_3	TTCCATTACCTGTCAGAACC
Nf2	Nf2_1	CACTGGGGCTTCGGGAAACC
Nf2	Nf2_2	TGGACTGCAGTATACAATCA
Nf2	Nf2_3	GTTGGATCATGATGTTTCGA
Notch2	Notch2_1	GCACCCTCGCCCTGCGTCAA
Notch2	Notch2_2	CATTACCTGTGAAGCCGACT
Notch2	Notch2_3	ATTGCCGCTGTCCCCCTCAG
Nprl3	Nprl3_1	TTGACAGCATATCTGCTCCG
Nprl3	Nprl3_2	GGAGCTTCCCGCTTTGGAGA
Nprl3	Nprl3_3	TACCACGTTAAAAAGAATCA
Nras	Nras_1	CTTCGCCTGTCCTCATGTAC
Nras	Nras_2	AAGCTCAACAGCAGTGACGA
Nras	Nras_3	CTGGGCCTCACCTCTATGGT
Pbrm1	Pbrm1_1	TAACACCATCCGAGACTATA
Pbrm1	Pbrm1_2	TTACAAGCAGCTTTATACTC
Pbrm1	Pbrm1_3	TCAACCAGACTATTATGAAG
Phf20l1	Phf20l1_1	GATGACTCCAGCGCTCAAAA
Phf20l1	Phf20l1_2	TTTGTTAATTGCTTCAATCT
Phf20l1	Phf20l1_3	GCACATTAAAGCCATGCCTG
Pik3ca	Pik3ca_1	TACTTCTGCTTGTCGTTGTT
Pik3ca	Pik3ca_2	GAATGTTCCTTCGAAAGTCT
Pik3ca	Pik3ca_3	TCTTACCTAGGATTGGAACA
Pik3r1	Pik3r1_1	TTGAGTTAATAAACCACTAC
Pik3r1	Pik3r1_2	GACGCATCTACTAAAATGCA
Pik3r1	Pik3r1_3	GTACCATTCAGCATCCTGCA
Plch2	Plch2_1	GACTACGGCTTGGATATGCC
Plch2	Plch2_2	AGCACCGCTCCTGCCTCCGA
Plch2	Plch2_3	ATTGGTCCGCTTCTACTACC
Plk2	Plk2_1	GAGTAGCTAAACCTCATCAG
Plk2	Plk2_2	GAAGGTGGGAGACTTTGGTT
Plk2	Plk2_3	CACCTCCTTCTGTGTTCCAG
Plxnb2	Plxnb2_1	TGCCCAGGAACGCAACAGTG
Plxnb2	Plxnb2_2	GCACCTCTCCTTGGGCCCGC
Plxnb2	Plxnb2_3	TTCCTCGGCCCGTGAGCACT
Rasa1	Rasa1_1	TATAGGATTATTGCTATGTG
Rasa1	Rasa1_2	CCTTATTGTTGAAGACCTAG
Rasa1	Rasa1_3	TACCTTATTTCATCAGTGTC
Rb1	Rb1_1	TCTTACCAGGATTCCATCCA
Rb1	Rb1_2	TGTAGCTCAGTAAAAGTGAA

Rb1	Rb1_3	TACAGAAAAGCATAGAAACC
Rbmx	Rbmx_1	CCCCTCCCGGGGAGGACACA
Rbmx	Rbmx_2	CTTTACCCATGTGTCCTCCC
Rbmx	Rbmx_3	GTTACGGAGGCCCGCCACGA
Rev1	Rev1_1	TCAGCTCTACACCAAGCCGT
Rev1	Rev1_2	TTTCTCAATTCCTCCGCAGA
Rev1	Rev1_3	CACCCTGTTGTTGAGCTGTT
Rexo4	Rexo4_1	ACTAGTATCTTCCTGTCCGT
Rexo4	Rexo4_2	TGTCTCATAGGGAAGACATC
Rexo4	Rexo4_3	AATGTTGAAGGGCAGAATCC
Rnf31	Rnf31_1	ATCCACCGTGCTGCGAAAGA
Rnf31	Rnf31_2	CAACGCTCTCTCTTAGCAGC
Rnf31	Rnf31_3	CTAAGAGAGAGCGTTGAAGA
Rnf43	Rnf43_1	TCCAGGCAAACTATGACGTG
Rnf43	Rnf43_2	GATGTCAAAGAGTACTGCGT
Rnf43	Rnf43_3	GTTTGTGTACAAGAATCGGA
Rps29	Rps29_1	CGCGTCTGCTCCAACCGCCA
Rps29	Rps29_2	TCACCCACGGAAGTTCGGCC
Rps29	Rps29_3	GGTGACTCCAGTAGAGCTGC
Sdk1	Sdk1_1	CGTTGCACCGTATTTCAAAA
Sdk1	Sdk1_2	TCTCTAAACCACGTCACTTG
Sdk1	Sdk1_3	GCCCCCAACCTCAAGTGACG
Setd2	Setd2_1	AGTTCCTCTCGATGTCCAAA
Setd2	Setd2_2	ATCGTCTCCTCATGATTGAA
Setd2	Setd2_3	CTGGGGCTTAAGGGCTGCTA
Sf3b1	Sf3b1_1	TTTGTATTCATCTTCCCGAT
Sf3b1	Sf3b1_2	TTCTCGCATAACATCCATGT
Sf3b1	Sf3b1_3	GCGAGAACAACATTTGACTA
Smad4	Smad4_1	AGACGGGCATAGATCACATG
Smad4	Smad4_2	CTGCCCGCCAGAATACTGGC
Smad4	Smad4_3	AGGTCAGCCGGCCAGTATTC
Smu1	Smu1_1	ATGACCATCGATTTGTTTCG
Smu1	Smu1_2	GAACAGTTTACATCGGGCAC
Smu1	Smu1_3	TCAAGTACTGCATGATTAGT
Snrpn	Snrpn_1	ATCTTCCTTACATCTTTAGG
Snrpn	Snrpn_2	GCTGGCAGAGGAGTGCCAGC
Snrpn	Snrpn_3	GAACTTGGTTTCAATGACTG
Stk11	Stk11_1	GTACTGCGTATGTGGCATGC
Stk11	Stk11_2	TCGGAATGTGATCCAGCTTG
Stk11	Stk11_3	TGACACCTGCCGGACAAGCC
Tbc1d8b	Tbc1d8b_1	GTGTACTTGTCAATTGCATG

Tbc1d8b	Tbc1d8b_2	AGCATTCACGTATGTTCTCA
Tbc1d8b	Tbc1d8b_3	TAATGAGACATACCTTCTTA
Tgfbr1	Tgfbr1_1	CATACAAATGGCCTGTCTCG
Tgfbr1	Tgfbr1_2	TTGACCTAATTCCTCGAGAC
Tgfbr1	Tgfbr1_3	GTTACGCCATGAAAATATCC
Thbs1	Thbs1_1	CACCATCGTGACCACTCTGC
Thbs1	Thbs1_2	ACTCCTCGTTGTTCTTGTAC
Thbs1	Thbs1_3	AGTCCAGTACAAGAACAACG
Tmtc1	Tmtc1_1	TTTCTAGGTGGCTGGCATCG
Tmtc1	Tmtc1_2	AAGAAGGGGGAAGCCGTTGT
Tmtc1	Tmtc1_3	AACCAAGCACACTCCAAACA
Tmtc4	Tmtc4_1	CTCCAGTCAGACACGCCCCT
Tmtc4	Tmtc4_2	GATACTCAGCAATACCCAGA
Tmtc4	Tmtc4_3	TTACTTACCAGCAGGTAACA
Trp53	Trp53_1	GGCCTTACCAGGATATCTTC
Trp53	Trp53_2	ATAAGCCTGAAAATGTCTCC
Trp53	Trp53_3	GTGCTGTGACTTCTTGTAGA
Ttc8	Ttc8_1	CTTCCAGATGAGCTAGTGAT
Ttc8	Ttc8_2	CCTCCTGTTTGGTTAGTGCC
Ttc8	Ttc8_3	GAGGACCGACTCAAGCTGTC
Wdr62	Wdr62_1	ATTGGGTGAGAAAGCCACAC
Wdr62	Wdr62_2	TTTCCCTAGGTGCGTGGTGG
Wdr62	Wdr62_3	GATGGGAAGTACATAGTGAC
Zfp317	Zfp317_1	TGATTCTTGTCAGCAGAGTT
Zfp317	Zfp317_2	AGCATCTGATGATGTAAGAT
Zfp317	Zfp317_3	GACATTTCAAGATATTGCTG
Zfp36l2	Zfp36l2_1	CAAGAAGTCGATATCGTAGA
Zfp36l2	Zfp36l2_2	GACCCACAACTTCGCCGCCG
Zfp36l2	Zfp36l2_3	GACCGCTCGTTCAGCGAGAA

Non-Targeting Controls (common between both libraries)	
sgRNA ID	sgRNA Spacer Sequence
NonTargetingControlGuideForMouse_0004	ACGTGTAAGGCGAACGCCTT
NonTargetingControlGuideForMouse_0006	CCGCGCCGTTAGGGAACGAG
NonTargetingControlGuideForMouse_0008	ACCCATCGGGTGCGATATGG
NonTargetingControlGuideForMouse_0011	TACAGTTATACGTCGCGGTG
NonTargetingControlGuideForMouse_0021	ACCAACGCTACGATCCCGGA
NonTargetingControlGuideForMouse_0024	GACCAACCTTACGGTAACTC
NonTargetingControlGuideForMouse_0027	TCCCGGGAGGTACGGTGTAC
NonTargetingControlGuideForMouse_0031	GTGGTTACGTTAACGACTAC
NonTargetingControlGuideForMouse_0040	GACCTCGCAATTGAGCGCTC
NonTargetingControlGuideForMouse_0046	GCGCGAGGGCACCGACAAGT
NonTargetingControlGuideForMouse_0059	TCAAACGCCCGGGCGCCCCA
NonTargetingControlGuideForMouse_0069	AGCTTAGATCGTGCGTCGTA
NonTargetingControlGuideForMouse_0075	AATGAAGCACCGATTGCGGA
NonTargetingControlGuideForMouse_0078	GGACTGAAACCGATAGTATC
NonTargetingControlGuideForMouse_0107	TAATGGAAGATGCGCGATAC
NonTargetingControlGuideForMouse_0112	CGAATGCGCCGGAGAATATT
NonTargetingControlGuideForMouse_0114	AGCCTAGTCGCGCTAATATT
NonTargetingControlGuideForMouse_0119	GGATAGCCCGGTTGGTGCGT
NonTargetingControlGuideForMouse_0127	TTATCGCACAACCCCGAAAG
NonTargetingControlGuideForMouse_0136	GTTTCCCGGGACTGTCGCGT
NonTargetingControlGuideForMouse_0142	ACGCTATAGTGTACGTCTAA
NonTargetingControlGuideForMouse_0149	AACCCCGGCTGTCATCGCCG
NonTargetingControlGuideForMouse_0150	GATTCATTATCGGCATACGG
NonTargetingControlGuideForMouse_0153	TTGCGTGTGCGTTGTTAACG
NonTargetingControlGuideForMouse_0158	ACCAGGACTGCCGCGTGAGG
NonTargetingControlGuideForMouse_0159	GCGCGTACATATAAATAGGT
NonTargetingControlGuideForMouse_0164	TAAACGATTCACCGATAACA
NonTargetingControlGuideForMouse_0171	TCTAGGATACTCTTAACGGG
NonTargetingControlGuideForMouse_0179	CTCGGCTTTACGATCGATCA
NonTargetingControlGuideForMouse_0183	CATTACGTGTCGAGCTCCGG
NonTargetingControlGuideForMouse_0190	CGTGCGTTCGTAATAAAAGG
NonTargetingControlGuideForMouse_0192	CGTGATTCCTAAGCCCCCGC
NonTargetingControlGuideForMouse_0194	GGCTTTACGTAAGGAGCGTA
NonTargetingControlGuideForMouse_0195	ACCCTAGCCTCATCGCGACC
NonTargetingControlGuideForMouse_0197	GTTACCCCTTTGGCCGGAAG
NonTargetingControlGuideForMouse_0200	CTGCACGGCACGACATCCAA
NonTargetingControlGuideForMouse_0203	GAGATTAAATTAACGCCGGC
NonTargetingControlGuideForMouse_0204	CGAATCGGGAAGGCGCGTGT
NonTargetingControlGuideForMouse_0208	TCCTGCGTTCCACTCGTACT

NonTargetingControlGuideForMouse_0213	TGTACTCGCATAGCGGGGGC
NonTargetingControlGuideForMouse_0217	TGCGGCAATGTTAACCCTTA
NonTargetingControlGuideForMouse_0220	GATAAGACTCCGCGAGCTTC
NonTargetingControlGuideForMouse_0225	TCGTAAGTTCGCTATATGCC
NonTargetingControlGuideForMouse_0230	AAACTCATACGTAGCGAATC
NonTargetingControlGuideForMouse_0231	CGATTGACGTTGGGCTCTCA
NonTargetingControlGuideForMouse_0238	CGCGCCTACCCTTTTACCGC
NonTargetingControlGuideForMouse_0243	GGAGTCTCACGCAATTAGCG
NonTargetingControlGuideForMouse_0248	GAGTCGAGTTAATAACGCTC
NonTargetingControlGuideForMouse_0254	CGTCACCGCTAGTAATGATG
NonTargetingControlGuideForMouse_0262	ACGGAACCCGATCGGAACGG
NonTargetingControlGuideForMouse_0264	ATTCCGTTTGCAGCGAGACC
NonTargetingControlGuideForMouse_0270	GTGATGGCCACGTCCGAACC
NonTargetingControlGuideForMouse_0271	TTGGACGTACACTTTCGTTC
NonTargetingControlGuideForMouse_0276	CGTTACGTTTCTTGCCAGGA
NonTargetingControlGuideForMouse_0277	TCTATATCTAGTCTCGGCGC
NonTargetingControlGuideForMouse_0285	TAGTTAACCGTAAAGTGGGC
NonTargetingControlGuideForMouse_0290	TTCGATATAGGGGACGGCGG
NonTargetingControlGuideForMouse_0295	TAATAAACTATGTCCCGCCG
NonTargetingControlGuideForMouse_0299	GCACCTCTAGCGCGCTCGGC
NonTargetingControlGuideForMouse_0303	GCCTATGTGAATCGCGAATT
NonTargetingControlGuideForMouse_0308	GATGGCCAGTAACGGCGTCA
NonTargetingControlGuideForMouse_0313	GGTTAAAAATTAAGCGGTCC
NonTargetingControlGuideForMouse_0318	CCCATCATTCGCGCTGACGT
NonTargetingControlGuideForMouse_0325	GGGCGGGACGTAATATTATG
NonTargetingControlGuideForMouse_0327	CTGATCGGTGCATATCTCGG
NonTargetingControlGuideForMouse_0331	GCTAAACGTATTTTACGGGC
NonTargetingControlGuideForMouse_0334	CTTTATTCCGTTGCATGTCG
NonTargetingControlGuideForMouse_0339	TCAACTTAACCTCGAGTCCG
NonTargetingControlGuideForMouse_0342	TGTAACGATCTGGGCGGTCT
NonTargetingControlGuideForMouse_0344	CAATATTCGACCTACGCTCC
NonTargetingControlGuideForMouse_0346	TGATTGGGGGTCGTTCGCCA
NonTargetingControlGuideForMouse_0347	TAGTCAGTCGGCCTCCGTGC
NonTargetingControlGuideForMouse_0349	GTGCCTGATAGTGTGAAGCG
NonTargetingControlGuideForMouse_0352	TTTTTAGACCTAATTCGCGC
NonTargetingControlGuideForMouse_0353	TCTACACGCGCGTTCAACCG
NonTargetingControlGuideForMouse_0359	CTCCGACGACTACGCAAGGA
NonTargetingControlGuideForMouse_0368	TATACTCGCCTGTCACAGCG
NonTargetingControlGuideForMouse_0373	CGTCAATGTGTCCGGACGGT
NonTargetingControlGuideForMouse_0379	TTCGACTGCGCACGCCATGA
NonTargetingControlGuideForMouse_0381	TGACTTGACACGTTCGATAT

NonTargetingControlGuideForMouse_0388	ACCCGCATATGCCGCCTAAG
NonTargetingControlGuideForMouse_0390	CACAATACGCGGCAGAGGTC
NonTargetingControlGuideForMouse_0394	AATTGTCTGATCGCGCCATA
NonTargetingControlGuideForMouse_0397	CTTATCGATTTGGGTTCAAC
NonTargetingControlGuideForMouse_0400	CAGCTGACCGTTAATCGATA
NonTargetingControlGuideForMouse_0409	TGTCGCCGATGGTCAGTCGC
NonTargetingControlGuideForMouse_0412	TCTCGAATAAATTTTCTCGC
NonTargetingControlGuideForMouse_0416	TTCCCCGCCCGTGCGGTCAT
NonTargetingControlGuideForMouse_0419	ATTATAGCAGCCCCCCGAAT
NonTargetingControlGuideForMouse_0424	GTTTCAGGAACGACGGCGAG
NonTargetingControlGuideForMouse_0429	TCGGCGCAGCCTAATGTATA
NonTargetingControlGuideForMouse_0432	GACAGTTGACGCGACGGAGA
NonTargetingControlGuideForMouse_0437	TGATGCCGGTACCCGTAACT
NonTargetingControlGuideForMouse_0440	TAATGCGAATGCGACCTCTC
NonTargetingControlGuideForMouse_0453	ATCATGCCTTCGCATTAACC
NonTargetingControlGuideForMouse_0458	CCCTTAAAGTGACGGACGAA
NonTargetingControlGuideForMouse_0462	TAGCCGAGTTCACGCCAGTA
NonTargetingControlGuideForMouse_0471	ATCGAGCACCGAGTTGTGAT
NonTargetingControlGuideForMouse_0473	CTACTTGTGACGACCTCGCG
NonTargetingControlGuideForMouse_0474	CTACGGATAGGCGCGGGTGA
NonTargetingControlGuideForMouse_0475	GCTTACCTACTCCGCCCCGC
NonTargetingControlGuideForMouse_0476	TAACACGCACTCACGTCCGG
NonTargetingControlGuideForMouse_0477	ACTCACCTCGCACGATCGTA
NonTargetingControlGuideForMouse_0478	ACACCCGTGTATGCACCGGG
NonTargetingControlGuideForMouse_0491	CGAGACGAATCCATCATGCG
NonTargetingControlGuideForMouse_0497	GTCCTAGATCCTATCGGGAG
NonTargetingControlGuideForMouse_0504	GTCGGCCGCTTAACCCTTTC
NonTargetingControlGuideForMouse_0509	AATAGGAACTCCGCACCCGA
NonTargetingControlGuideForMouse_0520	ATACGCGTGATGACCTTATG
NonTargetingControlGuideForMouse_0527	TTCCAGTATACCGAATTCGC
NonTargetingControlGuideForMouse_0535	CCGAGATGGCTCGGATAGAC
NonTargetingControlGuideForMouse_0542	GCCCGACCTCCACGTAAATC
NonTargetingControlGuideForMouse_0546	GCCGACCAACGATGACCACG
NonTargetingControlGuideForMouse_0558	GGTGTGGACCGCTTTTACGC
NonTargetingControlGuideForMouse_0564	CAATGGCGTCTGCCGTTCAG
NonTargetingControlGuideForMouse_0571	CACAACGCCTACCAGCGGAC
NonTargetingControlGuideForMouse_0580	TTCAAGTGTTATGGACGCGC
NonTargetingControlGuideForMouse_0588	GCTATTCCGCTCGTCAATTT
NonTargetingControlGuideForMouse_0593	ATGTGTATGAAGCCCGTCAT
NonTargetingControlGuideForMouse_0598	CGCGCGGATCTTCCGTACAA
NonTargetingControlGuideForMouse_0605	TCCCCTTCGTCGGCGCCAGG

NonTargetingControlGuideForMouse_0612	GACCGACGGTATACCCTACT
NonTargetingControlGuideForMouse_0618	GCCTTTTCCGCCCGTTCAAG
NonTargetingControlGuideForMouse_0622	CATTTCCGGGGTCCGATGCA
NonTargetingControlGuideForMouse_0625	CGAACTTCTACCGTTGTGCG
NonTargetingControlGuideForMouse_0629	CGTTATTGTACACCACGCTC
NonTargetingControlGuideForMouse_0630	CGAGCTCGGCAAATTTGATT
NonTargetingControlGuideForMouse_0635	CCTTTGATGGGCGCGTACTC
NonTargetingControlGuideForMouse_0640	GGAAACTCTTGCTCGACACG
NonTargetingControlGuideForMouse_0643	GGGCGCGTCATAGACACACG
NonTargetingControlGuideForMouse_0651	AGTCAGCTCCGGGGTATACG
NonTargetingControlGuideForMouse_0654	CATCGGTAGTCACTACGCCC
NonTargetingControlGuideForMouse_0657	CAAACGAGGCCACGTACGTG
NonTargetingControlGuideForMouse_0658	TCACTCAAAGTGCGTAGTCG
NonTargetingControlGuideForMouse_0667	AGTTAAGCGCGTATCTACGT
NonTargetingControlGuideForMouse_0668	AGTCCGGTCGAAATCTGTAT
NonTargetingControlGuideForMouse_0671	TGACCCGTCTCTTCGCGCAC
NonTargetingControlGuideForMouse_0678	TGCTGGCCGATACGTGCTTG
NonTargetingControlGuideForMouse_0687	GCCCGTTGTGAGCGGCATGC
NonTargetingControlGuideForMouse_0695	TGAGCCGGTGCGCTCTCGGT
NonTargetingControlGuideForMouse_0699	CGTTAATAAACGCGGTTTCA
NonTargetingControlGuideForMouse_0712	TAGAGCAGATCCCCTCAACG
NonTargetingControlGuideForMouse_0722	AAGCCTAACGGAGCTCGCGG
NonTargetingControlGuideForMouse_0723	ACCAATAACGGGTACAAGTC
NonTargetingControlGuideForMouse_0724	CACACGCCGCGGGCCGTGAG
NonTargetingControlGuideForMouse_0728	TTTTGGCGACGCATCGCTGG
NonTargetingControlGuideForMouse_0730	GACGTATTGGATTCGCAACC
NonTargetingControlGuideForMouse_0732	TTGAATCCGAGGCGCCGATG
NonTargetingControlGuideForMouse_0739	ACGAGGAGAAGTCGAGTATT
NonTargetingControlGuideForMouse_0740	GCTTCGACCACGGGTACTGC
NonTargetingControlGuideForMouse_0762	GTGTTGTCCGAGAACCCGAC
NonTargetingControlGuideForMouse_0765	GACTAGCACGAACTGCGGTT
NonTargetingControlGuideForMouse_0774	AGTCACTAGGCGCTCTCGCG
NonTargetingControlGuideForMouse_0775	CAGGTTTCCCCTAGTCGAAT
NonTargetingControlGuideForMouse_0780	CTTGACACCTCGTCCGCATC
NonTargetingControlGuideForMouse_0784	ATAATGCGCGCGAATACTTA
NonTargetingControlGuideForMouse_0788	ATGTTGCAGCGCTGTACGCC
NonTargetingControlGuideForMouse_0796	TCTTCTAGTGCGAGTTCCGC
NonTargetingControlGuideForMouse_0800	GAGCTAATCTGTAGACGTCG
NonTargetingControlGuideForMouse_0801	GCAGGGTTATTCCGCTTCTA
NonTargetingControlGuideForMouse_0803	CGTGTAGTCGGGTCGCATGT
NonTargetingControlGuideForMouse_0807	TGCCTACTGTCGTTCCCGCG

NonTargetingControlGuideForMouse_0808	GACCGGCCAACGGTAGCGGC
NonTargetingControlGuideForMouse_0809	ACAGAGTCGTGTGCCGAACG
NonTargetingControlGuideForMouse_0810	CTAAGTCGACTGTTTCGACC
NonTargetingControlGuideForMouse_0813	TCTCCGTCCGAAACGGCGAC
NonTargetingControlGuideForMouse_0817	AGTAGAGTCGCGAACGCTAC
NonTargetingControlGuideForMouse_0821	GACAATCGACACGCCGCTTC
NonTargetingControlGuideForMouse_0822	ATGGGAAAGCCCACGACAAT
NonTargetingControlGuideForMouse_0837	ACCTTGTGTTCGACGGTAGG
NonTargetingControlGuideForMouse_0841	CAACACGTGCGAGCGAACTG
NonTargetingControlGuideForMouse_0847	GCTCGAAAGTACGCTTCCTT
NonTargetingControlGuideForMouse_0850	AACGTTATAGCTTCGTCTCT
NonTargetingControlGuideForMouse_0859	GGGCAACGTCATCTGGCGAC
NonTargetingControlGuideForMouse_0865	TCAGTCAGTCGGCCCCCTGC
NonTargetingControlGuideForMouse_0866	AGTACAGCCGACGAGTGCGA
NonTargetingControlGuideForMouse_0869	TCCTACACCACCCCGGCCGT
NonTargetingControlGuideForMouse_0874	CGTTGTTTGGAGCCCGTCCA
NonTargetingControlGuideForMouse_0875	AAAGACGTGCATTCAGCGAG
NonTargetingControlGuideForMouse_0876	TAATTCTTTAGTACGTGCAA
NonTargetingControlGuideForMouse_0877	AGTCCCACGTTTGCCGTCAA
NonTargetingControlGuideForMouse_0878	GCCGTTGGAAAACTCCGGCC
NonTargetingControlGuideForMouse_0879	ATCGCTAATTTACGAAGGCG
NonTargetingControlGuideForMouse_0882	TGTTAATGCCAACGGGACAT
NonTargetingControlGuideForMouse_0887	TGTCTATGTTTGGCCACGCA
NonTargetingControlGuideForMouse_0889	TAGTAAAAAATTCGCCCTCC
NonTargetingControlGuideForMouse_0893	CGTTGCGAATCCACCCTATT
NonTargetingControlGuideForMouse_0896	GCAATAGTCGCACGGGTGAT
NonTargetingControlGuideForMouse_0898	TCAATCTGCGGTGACTCGTT
NonTargetingControlGuideForMouse_0903	CCCGGTCGGGGATACTAACT
NonTargetingControlGuideForMouse_0910	CCGACCGTGTGAGAGTACGT
NonTargetingControlGuideForMouse_0917	AAGCCTACTTCACCGGTCGG
NonTargetingControlGuideForMouse_0920	CCAACATCGGGGCGCATAAT
NonTargetingControlGuideForMouse_0922	AAGTCCCCTTCGGGAACTCC
NonTargetingControlGuideForMouse_0926	TCGGACTATAGCGTGAACGT
NonTargetingControlGuideForMouse_0930	CGTCATGGCAAGATCGATAA
NonTargetingControlGuideForMouse_0935	TTGTCGCTAACTCAGAATAG
NonTargetingControlGuideForMouse_0938	GATAGAACAACAGCGGTGCA
NonTargetingControlGuideForMouse_0942	TGCTTTACCGCGTTGGGTAA
NonTargetingControlGuideForMouse_0943	ATCTTAACCAGCGCATCCCG

Appendix 2: Differentially Expressed Genes Between Nf2-A and KRAS^{G12D}-Library Tumours

Upregulated (ordered by log_2 fold change)				
Gene ID	Gene Name	log_2 Fold Change	P-value	P(adj)
MGP_FVBNJ_G0017302	Cfd	4.384444214	4.30E-06	0.00017903
MGP_FVBNJ_G0000052	Ighv7-1	3.531523959	8.97E-05	0.00252514
MGP_FVBNJ_G0022490	Adipoq	3.441007774	0.00102384	0.01950631
MGP_FVBNJ_G0033495	Slc5a5	3.01705766	0.00010906	0.00297187
MGP_FVBNJ_G0028916	Rbp7	3.002657264	1.50E-06	6.89E-05
MGP_FVBNJ_G0022423	Tbx1	2.999890466	1.17E-05	0.0004379
MGP_FVBNJ_G0022875	Sim2	2.986600227	0.00151647	0.02665906
MGP_FVBNJ_G0030971	B4galnt3	2.948664467	0.00023502	0.00573179
MGP_FVBNJ_G0027812	Penk	2.861201955	0.00094071	0.01820267
MGP_FVBNJ_G0026033	Mpped2	2.726239209	1.39E-07	7.96E-06
MGP_FVBNJ_G0029302	Fgfbp1	2.677871058	0.00202403	0.03324805
MGP_FVBNJ_G0020798	Fam107a	2.622394741	0.00020526	0.00513558
MGP_FVBNJ_G0033916	Vat1l	2.543625305	0.00316314	0.04672599
MGP_FVBNJ_G0021464	Cldn10	2.477684012	0.00053777	0.01155983
MGP_FVBNJ_G0020769	Isl1	2.428196089	0.0018201	0.03078293
MGP_FVBNJ_G0023563	Adgrf4	2.414631081	0.00131364	0.02380684
MGP_FVBNJ_G0015259	Gm13429	2.295997381	0.00031079	0.00721642
MGP_FVBNJ_G0019067	Higd1b	2.287231062	0.00263376	0.04071374
MGP_FVBNJ_G0019350	Cbr2	2.240790946	2.72E-05	0.00092037
MGP_FVBNJ_G0016481	Pappa2	2.150865873	0.00240546	0.0382455
MGP_FVBNJ_G0003801	AI115009	2.062556039	0.00077162	0.01556578
MGP_FVBNJ_G0025147	Sfmbt2	2.060721612	0.00278649	0.0424563
MGP_FVBNJ_G0029441	Gnrhr	2.057138989	2.13E-05	0.00074896
MGP_FVBNJ_G0030245	Wnt2	2.036402468	0.00067847	0.01398462
MGP_FVBNJ_G0024052	Reep2	2.021313946	5.34E-06	0.00021694
MGP_FVBNJ_G0024951	Loxl4	2.003282231	0.0001197	0.00320698
MGP_FVBNJ_G0017158	Zfp365	1.991244019	7.31E-05	0.00213168
MGP_FVBNJ_G0030085	Papolb	1.969554203	0.00074618	0.01513477
MGP_FVBNJ_G0030996	Mfap5	1.943287185	0.00088864	0.01739127
MGP_FVBNJ_G0005388	1700018A04Rik	1.936628277	0.00311443	0.04619149
MGP_FVBNJ_G0017489	Usp44	1.925937622	0.00141853	0.02529642
MGP_FVBNJ_G0030084	Radil	1.912447317	0.00094001	0.01820267
MGP_FVBNJ_G0030963	Lrtm2	1.897101298	0.00052371	0.01133409

MGP_FVBNJ_G0027776	Ak5	1.865973689	0.00128694	0.02339901
MGP_FVBNJ_G0017931	Ccdc85a	1.855631533	6.60E-11	7.60E-09
MGP_FVBNJ_G0025605	Lrp2	1.835430232	0.00218612	0.03526101
MGP_FVBNJ_G0022142	Aqp2	1.756186625	0.00103984	0.01971004
MGP_FVBNJ_G0018124	Sowaha	1.688970804	0.00013522	0.00356299
MGP_FVBNJ_G0026018	Tcp11l1	1.684517242	0.00014091	0.00369534
MGP_FVBNJ_G0034571	Slc35f2	1.651483327	0.00252358	0.03946036
MGP_FVBNJ_G0019625	Trim9	1.648437137	0.00312055	0.04625017
MGP_FVBNJ_G0016854	Plagl1	1.605543295	0.00238969	0.03803878
MGP_FVBNJ_G0029605	Ephx4	1.603218216	5.79E-05	0.00174534
MGP_FVBNJ_G0025567	Pla2r1	1.595036746	6.63E-05	0.00195846
MGP_FVBNJ_G0002865	Gm12002	1.507153422	0.00187614	0.0314208
MGP_FVBNJ_G0023941	Greb1l	1.474769316	1.25E-06	5.87E-05
MGP_FVBNJ_G0018072	NA	1.4175676	0.00088828	0.01739127
MGP_FVBNJ_G0004509	0610038B21Rik	1.41726054	0.00062948	0.01312668
MGP_FVBNJ_G0026541	Procr	1.38226708	0.00344848	0.04994811
MGP_FVBNJ_G0036701	Gm10855	1.356292315	0.00258871	0.04021253
MGP_FVBNJ_G0025438	Ralgps1	1.349843877	3.11E-05	0.00103608
MGP_FVBNJ_G0015847	Sdpr	1.315638979	0.00161064	0.0278532
MGP_FVBNJ_G0025582	Cobll1	1.300943926	0.00196631	0.03251561
MGP_FVBNJ_G0026153	Jmjd7	1.27181428	1.72E-05	0.00062116
MGP_FVBNJ_G0008900	Mir5620	1.261654801	0.00328314	0.04802125
MGP_FVBNJ_G0025930	Agbl2	1.242094136	7.56E-06	0.00029732
MGP_FVBNJ_G0007022	Lockd	1.24094661	0.00218293	0.03526059
MGP_FVBNJ_G0033563	Arhgap10	1.226931229	0.00132863	0.02401991
MGP_FVBNJ_G0023062	Hcfc1r1	1.22689975	0.00022094	0.00544208
MGP_FVBNJ_G0001956	NA	1.223950897	0.00236728	0.03771908
MGP_FVBNJ_G0015302	Gm13986	1.212166039	0.00127813	0.02327684
MGP_FVBNJ_G0025373	Sh3glb2	1.202704656	0.00010798	0.00295331
MGP_FVBNJ_G0034950	Amotl2	1.197275996	0.00066371	0.01376252
MGP_FVBNJ_G0035976	Gja6	1.189200704	0.00039444	0.00888132
MGP_FVBNJ_G0035816	Drp2	1.1644393	0.00210762	0.03439314
MGP_FVBNJ_G0026143	Oip5	1.163334982	0.00170497	0.02911785
MGP_FVBNJ_G0017810	Gatsl3	1.149881813	0.00120071	0.02221161
MGP_FVBNJ_G0027560	Sort1	1.142919002	0.00191276	0.03193802
MGP_FVBNJ_G0039038	Gm23995	1.100300736	0.00329901	0.04815857
MGP_FVBNJ_G0001698	NA	1.09479016	0.00184461	0.03107951
MGP_FVBNJ_G0017312	Abca7	1.080352743	0.00292913	0.04414696
MGP_FVBNJ_G0027188	NA	1.077752708	0.00176564	0.0299756
MGP_FVBNJ_G0023905	Ston1	1.071408489	0.00034458	0.00790157
MGP_FVBNJ_G0023605	Tjap1	1.056027424	0.00156015	0.02727627

MGP_FVBNJ_G0026093	Fmn1	1.052513083	9.88E-08	5.81E-06
MGP_FVBNJ_G0023829	NA	1.022093848	0.00028916	0.00678495
MGP_FVBNJ_G0035173	Oxsr1	1.008795814	8.49E-06	0.00032989
MGP_FVBNJ_G0018105	Jade2	0.986703817	5.79E-05	0.00174534
MGP_FVBNJ_G0024625	Rom1	0.98451075	0.00097129	0.01866466
MGP_FVBNJ_G0022529	Apod	0.962329671	0.00292914	0.04414696
MGP_FVBNJ_G0027753	Ssx2ip	0.961047584	0.00219437	0.03531745
MGP_FVBNJ_G0034889	Plscr2	0.932352376	8.74E-06	0.00033655
MGP_FVBNJ_G0029611	Glmn	0.919366654	4.08E-05	0.00130435
MGP_FVBNJ_G0035334	Maoa	0.913216001	0.00087218	0.01717484
MGP_FVBNJ_G0034031	2310022B05Rik	0.912224779	0.00170459	0.02911785
MGP_FVBNJ_G0025627	Slc25a12	0.90788819	3.91E-05	0.00126163
MGP_FVBNJ_G0039251	Gm24299	0.902839194	0.00187604	0.0314208
MGP_FVBNJ_G0029425	Paics	0.896507534	0.00027679	0.0065222
MGP_FVBNJ_G0032777	Fam57b	0.896403058	0.00150953	0.02655792
MGP_FVBNJ_G0023901	Msh6	0.892236189	0.00100998	0.01930838
MGP_FVBNJ_G0025320	Vav2	0.883717835	2.29E-06	0.00010049
MGP_FVBNJ_G0016860	Adat2	0.878729829	0.00242678	0.03847459
MGP_FVBNJ_G0035972	Phka2	0.878709039	0.00091822	0.01784475
MGP_FVBNJ_G0018106	Cdkn2aipnl	0.874432937	0.00269283	0.04131175
MGP_FVBNJ_G0034560	Zc3h12c	0.867391051	0.00061858	0.01298345
MGP_FVBNJ_G0035975	Cdkl5	0.858156769	0.00268451	0.04128349
MGP_FVBNJ_G0029882	Bri3bp	0.84613653	0.00317342	0.04684689
MGP_FVBNJ_G0035190	Rpsa	0.833780103	0.00245211	0.03868362
MGP_FVBNJ_G0029616	Rpl5	0.829440347	0.00218503	0.03526101
MGP_FVBNJ_G0023202	Ergic1	0.824471715	0.00268186	0.04128349
MGP_FVBNJ_G0023623	Ppp2r5d	0.824109067	0.00104821	0.01985185
MGP_FVBNJ_G0034927	Pik3cb	0.820554883	5.53E-09	3.97E-07
MGP_FVBNJ_G0012059	NA	0.811105878	0.00079407	0.01588927
MGP_FVBNJ_G0013118	NA	0.796599148	0.00057979	0.01223875
MGP_FVBNJ_G0025108	Nmt2	0.793064992	0.00027534	0.00649726
MGP_FVBNJ_G0033659	Zfp791	0.79009974	0.00130963	0.02375343
MGP_FVBNJ_G0026825	Zfp512b	0.781702475	0.00101593	0.01940536
MGP_FVBNJ_G0023647	Usp49	0.778210186	0.00299379	0.04479223
MGP_FVBNJ_G0034691	Rpl4	0.766892801	8.66E-07	4.17E-05
MGP_FVBNJ_G0025292	Snapc4	0.760445878	0.00126311	0.02307753
MGP_FVBNJ_G0024654	Ddb1	0.75990353	0.00243538	0.03857548
MGP_FVBNJ_G0017433	Mterf2	0.756723539	0.0014323	0.02546041
MGP_FVBNJ_G0026034	Arl14ep	0.755478306	7.36E-05	0.0021428
MGP_FVBNJ_G0023118	NA	0.752363827	0.00265336	0.04091411
MGP_FVBNJ_G0025559	Pkp4	0.740499929	0.0007449	0.01513477

MGP_FVBNJ_G0033365	Wwc2	0.737010351	0.00091664	0.01782966
MGP_FVBNJ_G0025639	Ola1	0.73367698	0.00246368	0.03883862
MGP_FVBNJ_G0029617	Fam69a	0.731711548	0.00032001	0.00739975
MGP_FVBNJ_G0033072	Arhgef18	0.721482557	0.00096964	0.01865142
MGP_FVBNJ_G0034122	Taf1d	0.721128546	0.00078669	0.01579737
MGP_FVBNJ_G0029016	Fam132a	0.719086343	0.00021622	0.00533739
MGP_FVBNJ_G0010183	NA	0.717074762	0.00062951	0.01312668
MGP_FVBNJ_G0034559	Rdx	0.716915204	0.0020956	0.03424704
MGP_FVBNJ_G0029860	Mphosph9	0.71525807	0.00013886	0.00365105
MGP_FVBNJ_G0030242	Met	0.712943991	0.00055663	0.01185082
MGP_FVBNJ_G0024591	Rtn3	0.701470375	0.00096263	0.01854625
MGP_FVBNJ_G0029625	Mfsd7a	0.700534177	0.00206663	0.03382317
MGP_FVBNJ_G0026271	Anapc1	0.696206296	1.16E-09	9.83E-08
MGP_FVBNJ_G0030049	Sun1	0.693894902	0.00067058	0.0138922
MGP_FVBNJ_G0030124	Trrap	0.69102864	0.00067229	0.01390178
MGP_FVBNJ_G0009909	NA	0.689469188	1.83E-05	0.00065867
MGP_FVBNJ_G0034564	Kdelc2	0.685701043	0.00020226	0.00508867
MGP_FVBNJ_G0033012	Ap2a2	0.680958219	4.22E-05	0.00133772
MGP_FVBNJ_G0030147	Cdk8	0.680325081	0.00047279	0.0103833
MGP_FVBNJ_G0026823	Dnajc5	0.679334155	0.0027987	0.04255507
MGP_FVBNJ_G0026546	Uqcc1	0.669294378	0.00102977	0.01958587
MGP_FVBNJ_G0025306	Rpl7a	0.668802005	0.0019721	0.03258713
MGP_FVBNJ_G0024882	Ide	0.66616497	0.00164492	0.02837999
MGP_FVBNJ_G0026366	Ndufaf5	0.653415171	0.0015337	0.02689833
MGP_FVBNJ_G0026279	Ttl	0.653392607	0.00301409	0.04506014
MGP_FVBNJ_G0011290	NA	0.653235547	0.00013317	0.00352147
MGP_FVBNJ_G0030187	Stard13	0.651863109	0.00185085	0.03116112
MGP_FVBNJ_G0021902	Rpl3	0.647110428	0.00074309	0.01512716
MGP_FVBNJ_G0025137	Usp6nl	0.644195604	0.00030806	0.00716062
MGP_FVBNJ_G0029298	Cc2d2a	0.638135443	0.00217295	0.03512492
MGP_FVBNJ_G0029812	Ppp1cc	0.637435569	6.09E-05	0.00182272
MGP_FVBNJ_G0025157	Fam188a	0.631150315	0.00045236	0.00998162
MGP_FVBNJ_G0026040	Lin7c	0.63031777	2.76E-05	0.00093352
MGP_FVBNJ_G0026024	Eif3m	0.62904358	0.00244477	0.03859511
MGP_FVBNJ_G0024439	Doc2g	0.62659339	0.00136577	0.02443364
MGP_FVBNJ_G0025193	Arhgap21	0.623397842	0.00124999	0.02289525
MGP_FVBNJ_G0027135	Mnd1	0.615260215	0.0017552	0.02984376
MGP_FVBNJ_G0035203	Sec22c	0.614931941	0.00134296	0.02416136
MGP_FVBNJ_G0029914	Sbds	0.610638813	0.00067217	0.01390178
MGP_FVBNJ_G0025052	Smc3	0.609019458	0.0018899	0.03157998
MGP_FVBNJ_G0024588	Mark2	0.608772467	0.00066089	0.01371684

MGP_FVBNJ_G0034567	Npat	0.606336038	0.00282943	0.04293455
MGP_FVBNJ_G0026149	Rpap1	0.605516184	0.00042763	0.00951332
MGP_FVBNJ_G0024765	Vps13a	0.603877458	0.00157287	0.02741279
MGP_FVBNJ_G0024422	Mrpl21	0.598308729	0.00259982	0.04030081
MGP_FVBNJ_G0022960	Tcp1	0.597753734	1.19E-05	0.00044529
MGP_FVBNJ_G0023588	Hsp90ab1	0.597048078	0.00027154	0.00641876
MGP_FVBNJ_G0029646	NA	0.593728207	8.12E-05	0.00233086
MGP_FVBNJ_G0029827	Anapc5	0.59219497	1.11E-05	0.00042043
MGP_FVBNJ_G0026126	Rpusd2	0.591733469	0.00126624	0.02307918
MGP_FVBNJ_G0023903	Foxn2	0.583834283	4.37E-15	1.43E-12
MGP_FVBNJ_G0026408	Napb	0.581997262	0.00079404	0.01588927
MGP_FVBNJ_G0024785	Abhd17b	0.581228467	0.00166162	0.02853563
MGP_FVBNJ_G0023834	NA	0.579682275	0.00165819	0.02853563
MGP_FVBNJ_G0025222	Dph7	0.575599157	0.00088565	0.01736319
MGP_FVBNJ_G0025517	Scai	0.574853487	0.00156393	0.02732096
MGP_FVBNJ_G0026822	Tpd52l2	0.570143962	0.00158348	0.02757598
MGP_FVBNJ_G0029870	Eif2b1	0.564745181	0.00077868	0.01569397
MGP_FVBNJ_G0026084	Slc12a6	0.563201026	8.81E-05	0.00248646
MGP_FVBNJ_G0029198	Zfp512	0.562930006	0.00286992	0.0434307
MGP_FVBNJ_G0029794	Gm15800	0.554106309	0.00091184	0.0177829
MGP_FVBNJ_G0023601	Xpo5	0.554004179	0.00230036	0.03678424
MGP_FVBNJ_G0023741	Gtf2f1	0.553542609	6.54E-05	0.00193781
MGP_FVBNJ_G0026165	Zfp106	0.552477794	0.00165092	0.02846148
MGP_FVBNJ_G0026560	Phf20	0.550065753	0.000476	0.01044337
MGP_FVBNJ_G0025569	Rbms1	0.545716723	0.00309477	0.04602103
MGP_FVBNJ_G0026101	Dph6	0.544499688	0.00032695	0.00753679
MGP_FVBNJ_G0029622	Pigg	0.543235974	0.00327354	0.04794379
MGP_FVBNJ_G0025265	Rabl6	0.530696007	0.00014115	0.00369734
MGP_FVBNJ_G0029801	Mapkapk5	0.5301103	2.55E-07	1.37E-05
MGP_FVBNJ_G0026249	Snrnp200	0.529738393	5.24E-05	0.00161017
MGP_FVBNJ_G0015655	NA	0.521927404	0.00053185	0.01145449
MGP_FVBNJ_G0036001	Mospd2	0.518860885	0.00244248	0.03858639
MGP_FVBNJ_G0026384	Polr3f	0.510819105	0.00062284	0.01304861
MGP_FVBNJ_G0030101	Zdhhc4	0.508874228	0.00262799	0.04065316
MGP_FVBNJ_G0012583	NA	0.508553191	0.00319556	0.0471003
MGP_FVBNJ_G0026232	Galk2	0.505728048	0.00062671	0.01311747
MGP_FVBNJ_G0026021	Qser1	0.49355029	0.00041507	0.00928016
MGP_FVBNJ_G0029629	Tmem175	0.486061524	0.00174032	0.02963599
MGP_FVBNJ_G0025622	Tlk1	0.483560653	0.00145218	0.02575228
MGP_FVBNJ_G0034522	Usp28	0.477528667	5.90E-05	0.0017759
MGP_FVBNJ_G0027030	Eif2a	0.475539742	0.00113344	0.02114277

MGP_FVBNJ_G0026303	Ptpra	0.473132197	0.00184328	0.03107951
MGP_FVBNJ_G0025150	Rbm17	0.470488072	0.00330987	0.04828549
MGP_FVBNJ_G0035229	Zdhhc3	0.460064002	0.00017589	0.00449638
MGP_FVBNJ_G0025154	Fbxo18	0.455143407	1.89E-05	0.0006739
MGP_FVBNJ_G0022906	NA	0.452879015	0.00241962	0.03840706
MGP_FVBNJ_G0026398	Xrn2	0.451313861	0.00011798	0.0031762
MGP_FVBNJ_G0023270	Cmtr1	0.447912087	0.00187561	0.0314208
MGP_FVBNJ_G0026600	Dhx35	0.447836774	0.00083957	0.01661223
MGP_FVBNJ_G0025295	Inpp5e	0.443147988	0.00050123	0.010911
MGP_FVBNJ_G0026514	Cbfa2t2	0.440252979	0.00054033	0.01159246
MGP_FVBNJ_G0022432	Hira	0.435705897	0.00045192	0.00998162
MGP_FVBNJ_G0023823	Birc6	0.434568175	0.0015929	0.02765368
MGP_FVBNJ_G0029554	Hnrnpdl	0.426952598	0.00113784	0.02120694
MGP_FVBNJ_G0029575	Aff1	0.420130652	0.00069263	0.0142429
MGP_FVBNJ_G0025609	Ppig	0.414584816	0.00326827	0.04792959
MGP_FVBNJ_G0026538	Trpc4ap	0.411499828	1.82E-05	0.00065404
MGP_FVBNJ_G0031449	NA	0.407922678	2.83E-05	0.00095127
MGP_FVBNJ_G0024789	Smc5	0.405854248	0.00256676	0.03995339
MGP_FVBNJ_G0025499	Rc3h2	0.397804641	0.00100667	0.01927457
MGP_FVBNJ_G0029892	Stx2	0.382323434	0.00256089	0.03989153
MGP_FVBNJ_G0024608	Stx5a	0.379879534	0.00126554	0.02307918
MGP_FVBNJ_G0026826	Prpf6	0.375467124	0.00121409	0.02240472
MGP_FVBNJ_G0025446	Gapvd1	0.37542879	0.00179021	0.03032337
MGP_FVBNJ_G0024051	Kdm3b	0.364493354	0.00136382	0.02441845
MGP_FVBNJ_G0025180	Mllt10	0.356072343	8.37E-05	0.00239088
MGP_FVBNJ_G0026533	Ncoa6	0.346287513	0.00021251	0.00526914
MGP_FVBNJ_G0024233	Snx2	0.342697666	0.00319693	0.0471003
MGP_FVBNJ_G0026491	Asxl1	0.307385519	0.0014512	0.02575228
MGP_FVBNJ_G0029844	Vps33a	0.281174352	0.00031618	0.00731881
MGP_FVBNJ_G0022963	NA	0.24321458	0.00149198	0.02631819
MGP_FVBNJ_G0023280	Tff3	0.206571169	8.37E-11	9.37E-09

Downregulated (ordered by log_2 fold change)				
Gene ID	Gene Name	log_2 Fold Change	P-value	P(adj)
MGP_FVBNJ_G0033334	Klkb1	-6.2311663	1.60E-16	8.09E-14
MGP_FVBNJ_G0033786	Ces3b	-6.1677257	2.37E-24	4.31E-20
MGP_FVBNJ_G0039773	NA	-5.8351415	1.75E-18	1.77E-15
MGP_FVBNJ_G0027127	Fgg	-5.6133256	2.08E-21	1.55E-17
MGP_FVBNJ_G0020918	Itih4	-5.6096541	3.87E-24	4.31E-20
MGP_FVBNJ_G0030990	A2m	-5.5528682	2.67E-20	7.44E-17
MGP_FVBNJ_G0039590	NA	-5.4370853	2.34E-19	4.01E-16
MGP_FVBNJ_G0027129	Fgb	-5.4211464	2.08E-20	6.62E-17
MGP_FVBNJ_G0027128	Fga	-5.4098092	1.64E-19	3.05E-16
MGP_FVBNJ_G0021006	Mbl1	-5.4071587	8.42E-19	9.88E-16
MGP_FVBNJ_G0028291	C8b	-5.358267	4.10E-20	9.53E-17
MGP_FVBNJ_G0028292	C8a	-5.3397157	4.96E-18	3.95E-15
MGP_FVBNJ_G0039589	NA	-5.3010204	4.16E-18	3.45E-15
MGP_FVBNJ_G0017676	Hsd17b6	-5.2820894	2.18E-14	5.51E-12
MGP_FVBNJ_G0019870	Serpina10	-5.2721687	6.98E-21	3.89E-17
MGP_FVBNJ_G0028146	Orm2	-5.2119906	3.35E-19	5.17E-16
MGP_FVBNJ_G0020920	Itih1	-5.2027968	7.98E-19	9.88E-16
MGP_FVBNJ_G0016639	Apcs	-5.1702064	1.28E-17	9.82E-15
MGP_FVBNJ_G0018533	Serpinf2	-5.1688434	4.28E-20	9.53E-17
MGP_FVBNJ_G0035319	Otc	-5.1427099	4.67E-16	2.04E-13
MGP_FVBNJ_G0024913	Cyp2c70	-5.1406436	3.51E-19	5.17E-16
MGP_FVBNJ_G0019875	NA	-5.0743681	3.71E-19	5.17E-16
MGP_FVBNJ_G0032049	Saa4	-5.0416173	3.16E-17	2.07E-14
MGP_FVBNJ_G0025065	Habp2	-5.0114251	1.43E-18	1.52E-15
MGP_FVBNJ_G0020919	Itih3	-4.9936076	5.09E-16	2.18E-13
MGP_FVBNJ_G0023309	Cyp4f15	-4.9803253	4.88E-15	1.53E-12
MGP_FVBNJ_G0031942	Aspdh	-4.9798737	1.16E-12	1.95E-10
MGP_FVBNJ_G0022485	Fetub	-4.9638824	1.95E-20	6.62E-17
MGP_FVBNJ_G0021528	Ugt3a2	-4.9254362	8.79E-17	4.90E-14
MGP_FVBNJ_G0020185	Slc17a3	-4.925287	1.23E-11	1.66E-09
MGP_FVBNJ_G0019874	NA	-4.9252644	2.25E-15	8.08E-13
MGP_FVBNJ_G0016401	F13b	-4.9144832	7.09E-17	4.16E-14
MGP_FVBNJ_G0028142	Ambp	-4.9122599	4.18E-18	3.45E-15
MGP_FVBNJ_G0032012	Fgf21	-4.8996402	5.29E-11	6.30E-09
MGP_FVBNJ_G0018604	Vtn	-4.8819821	2.23E-17	1.55E-14
MGP_FVBNJ_G0032531	Hpx	-4.8496338	1.57E-17	1.17E-14
MGP_FVBNJ_G0024597	Slc22a26	-4.8485013	3.02E-15	1.02E-12

MGP_FVBNJ_G0022519	Cpn2	-4.828484	6.19E-17	3.73E-14
MGP_FVBNJ_G0024601	NA	-4.8237556	4.74E-10	4.37E-08
MGP_FVBNJ_G0016084	2810459M11Rik	-4.8198836	5.35E-17	3.36E-14
MGP_FVBNJ_G0029496	5830473C10Rik	-4.8129373	5.08E-11	6.09E-09
MGP_FVBNJ_G0023422	Apom	-4.800059	2.59E-16	1.23E-13
MGP_FVBNJ_G0024907	NA	-4.7783821	5.69E-12	8.19E-10
MGP_FVBNJ_G0029632	Slc26a1	-4.7735507	9.49E-17	5.16E-14
MGP_FVBNJ_G0017654	Inhbc	-4.7505073	2.16E-14	5.51E-12
MGP_FVBNJ_G0021962	Cyp2d26	-4.743761	1.21E-16	6.45E-14
MGP_FVBNJ_G0029495	Afm	-4.7301546	5.61E-16	2.36E-13
MGP_FVBNJ_G0033324	Fgl1	-4.725988	4.38E-15	1.43E-12
MGP_FVBNJ_G0031097	Pzp	-4.7243273	7.11E-15	2.06E-12
MGP_FVBNJ_G0028061	Plppr1	-4.7213843	2.77E-10	2.71E-08
MGP_FVBNJ_G0032994	Cdhr5	-4.721059	2.27E-18	2.11E-15
MGP_FVBNJ_G0031829	Hamp	-4.7198636	5.42E-17	3.36E-14
MGP_FVBNJ_G0022951	Plg	-4.7146854	1.46E-15	5.62E-13
MGP_FVBNJ_G0024964	Cpn1	-4.6847047	7.61E-19	9.88E-16
MGP_FVBNJ_G0032974	Nlrp6	-4.674676	1.24E-16	6.45E-14
MGP_FVBNJ_G0030204	Pon1	-4.6676627	3.19E-16	1.48E-13
MGP_FVBNJ_G0021379	Cpb2	-4.6661794	1.27E-14	3.49E-12
MGP_FVBNJ_G0022487	Kng2	-4.6506519	3.00E-18	2.67E-15
MGP_FVBNJ_G0019876	Serpina11	-4.6378938	2.97E-14	7.35E-12
MGP_FVBNJ_G0027758	Uox	-4.6332986	1.86E-15	6.91E-13
MGP_FVBNJ_G0021511	C9	-4.6291079	4.26E-16	1.90E-13
MGP_FVBNJ_G0035703	Gjb1	-4.6216713	1.96E-16	9.71E-14
MGP_FVBNJ_G0016638	Crp	-4.6100415	2.78E-15	9.54E-13
MGP_FVBNJ_G0031225	Slco1b2	-4.5709193	3.26E-16	1.49E-13
MGP_FVBNJ_G0019886	Serpina3k	-4.5678837	4.37E-14	1.04E-11
MGP_FVBNJ_G0024965	Cyp2c44	-4.5661113	4.76E-15	1.51E-12
MGP_FVBNJ_G0017675	Sdr9c7	-4.5617367	2.69E-10	2.66E-08
MGP_FVBNJ_G0022486	Hrg	-4.5546041	3.32E-14	8.13E-12
MGP_FVBNJ_G0035505	F9	-4.5534953	8.34E-17	4.77E-14
MGP_FVBNJ_G0025226	Entpd8	-4.5443081	1.45E-14	3.93E-12
MGP_FVBNJ_G0019149	Apoh	-4.5361295	1.20E-14	3.34E-12
MGP_FVBNJ_G0017653	Inhbe	-4.5259488	5.71E-16	2.36E-13
MGP_FVBNJ_G0024841	A1cf	-4.5230639	1.13E-14	3.19E-12
MGP_FVBNJ_G0033703	Ces1b	-4.5174264	5.86E-14	1.32E-11
MGP_FVBNJ_G0030021	Azgp1	-4.5157557	9.81E-14	1.95E-11
MGP_FVBNJ_G0029836	Hpd	-4.5127122	5.26E-15	1.61E-12
MGP_FVBNJ_G0030139	Cyp3a11	-4.5077488	2.03E-15	7.41E-13
MGP_FVBNJ_G0027126	Lrat	-4.5047112	2.28E-15	8.08E-13

MGP_FVBNJ_G0027659	Cfi	-4.5019598	1.12E-20	5.00E-17
MGP_FVBNJ_G0016178	Agxt	-4.4983026	5.99E-14	1.33E-11
MGP_FVBNJ_G0035026	Slc38a3	-4.4980716	7.27E-15	2.08E-12
MGP_FVBNJ_G0022874	Cldn14	-4.4933211	2.35E-11	3.04E-09
MGP_FVBNJ_G0022484	Ahsg	-4.4900382	4.70E-14	1.09E-11
MGP_FVBNJ_G0020478	Fbp1	-4.4855256	9.46E-16	3.77E-13
MGP_FVBNJ_G0030736	Nat8f2	-4.4836671	5.32E-15	1.61E-12
MGP_FVBNJ_G0027046	Aadac	-4.4781837	6.51E-14	1.41E-11
MGP_FVBNJ_G0031525	Foxa3	-4.4432345	4.26E-14	1.02E-11
MGP_FVBNJ_G0024600	NA	-4.4390621	3.95E-09	2.97E-07
MGP_FVBNJ_G0033333	F11	-4.4349639	5.17E-13	9.29E-11
MGP_FVBNJ_G0017687	Apon	-4.425947	2.12E-18	2.05E-15
MGP_FVBNJ_G0029487	Gc	-4.4255515	2.53E-17	1.71E-14
MGP_FVBNJ_G0018915	Krt12	-4.4223068	0.00011239	0.00304881
MGP_FVBNJ_G0029455	NA	-4.4206889	5.15E-09	3.77E-07
MGP_FVBNJ_G0028062	Acnat1	-4.4172389	2.10E-17	1.51E-14
MGP_FVBNJ_G0035863	Serpina7	-4.4115252	7.11E-12	1.01E-09
MGP_FVBNJ_G0033108	Slc10a2	-4.4029384	2.60E-12	4.08E-10
MGP_FVBNJ_G0017673	BC089597	-4.3969922	6.03E-14	1.33E-11
MGP_FVBNJ_G0027712	Adh4	-4.3957061	3.50E-13	6.35E-11
MGP_FVBNJ_G0020020	Akr1c20	-4.3903673	4.73E-12	7.22E-10
MGP_FVBNJ_G0024716	Gm4952	-4.3880438	1.20E-13	2.35E-11
MGP_FVBNJ_G0011167	NA	-4.3447887	1.38E-06	6.41E-05
MGP_FVBNJ_G0025950	F2	-4.3325294	5.94E-16	2.41E-13
MGP_FVBNJ_G0030143	NA	-4.3318253	7.65E-08	4.59E-06
MGP_FVBNJ_G0016490	Serpinc1	-4.330957	1.22E-15	4.78E-13
MGP_FVBNJ_G0016962	Sult3a1	-4.3299142	5.35E-15	1.61E-12
MGP_FVBNJ_G0023979	Ttr	-4.3207174	1.62E-14	4.34E-12
MGP_FVBNJ_G0021956	NA	-4.3162387	5.60E-11	6.53E-09
MGP_FVBNJ_G0026288	Pdyn	-4.3129267	4.38E-06	0.00018127
MGP_FVBNJ_G0031228	Slco1a1	-4.3114224	6.33E-14	1.38E-11
MGP_FVBNJ_G0030707	Clec4f	-4.30411	9.13E-13	1.55E-10
MGP_FVBNJ_G0021629	Dpys	-4.3014743	6.79E-14	1.45E-11
MGP_FVBNJ_G0017688	Apof	-4.2997185	1.65E-14	4.36E-12
MGP_FVBNJ_G0030808	Uroc1	-4.2952491	2.14E-13	4.04E-11
MGP_FVBNJ_G0024022	Myo7b	-4.2895389	1.16E-09	9.83E-08
MGP_FVBNJ_G0035212	Cyp8b1	-4.2854449	5.51E-12	8.14E-10
MGP_FVBNJ_G0016961	Sult3a2	-4.2828761	5.59E-12	8.14E-10
MGP_FVBNJ_G0026753	Pck1	-4.2775119	1.62E-13	3.11E-11
MGP_FVBNJ_G0032690	Acsm1	-4.271492	2.84E-13	5.23E-11
MGP_FVBNJ_G0033139	Proz	-4.267821	4.23E-14	1.02E-11

MGP_FVBNJ_G0021659	Colec10	-4.2647755	8.58E-12	1.19E-09
MGP_FVBNJ_G0027610	Dpyd	-4.2635306	1.02E-12	1.73E-10
MGP_FVBNJ_G0034897	Paqr9	-4.2564204	1.91E-14	4.96E-12
MGP_FVBNJ_G0028835	Agmat	-4.2543781	1.91E-13	3.64E-11
MGP_FVBNJ_G0016897	Aldh8a1	-4.2415175	2.32E-14	5.81E-12
MGP_FVBNJ_G0031201	Art4	-4.2383138	8.79E-14	1.81E-11
MGP_FVBNJ_G0028582	Tlr12	-4.2357345	2.10E-09	1.69E-07
MGP_FVBNJ_G0028258	Angptl3	-4.2341292	2.62E-11	3.35E-09
MGP_FVBNJ_G0030642	Sftpb	-4.2114463	0.00025069	0.00602817
MGP_FVBNJ_G0020184	Slc17a2	-4.206245	1.37E-13	2.66E-11
MGP_FVBNJ_G0029453	Ugt2b36	-4.2062355	5.52E-14	1.26E-11
MGP_FVBNJ_G0020635	NA	-4.2013658	4.74E-10	4.37E-08
MGP_FVBNJ_G0025454	NA	-4.1994694	1.93E-12	3.12E-10
MGP_FVBNJ_G0018916	Krt20	-4.178719	6.82E-07	3.35E-05
MGP_FVBNJ_G0023304	Pglyrp2	-4.1704282	6.24E-15	1.83E-12
MGP_FVBNJ_G0017674	Rdh7	-4.1577502	2.76E-12	4.27E-10
MGP_FVBNJ_G0020297	Elovl2	-4.1573164	1.70E-12	2.77E-10
MGP_FVBNJ_G0027759	Dnase2b	-4.1491404	2.80E-09	2.19E-07
MGP_FVBNJ_G0031069	Ntf3	-4.1435512	5.97E-10	5.39E-08
MGP_FVBNJ_G0021959	NA	-4.1267726	4.88E-12	7.40E-10
MGP_FVBNJ_G0026693	Slc13a3	-4.11711	5.25E-12	7.85E-10
MGP_FVBNJ_G0022954	Slc22a1	-4.1133405	1.21E-12	2.01E-10
MGP_FVBNJ_G0018425	Gltpd2	-4.1090909	6.96E-14	1.48E-11
MGP_FVBNJ_G0024905	Cyp2c39	-4.1013704	1.28E-12	2.09E-10
MGP_FVBNJ_G0034029	Agt	-4.0996019	1.66E-14	4.36E-12
MGP_FVBNJ_G0018406	Asgr2	-4.0984656	7.76E-14	1.62E-11
MGP_FVBNJ_G0033088	Clec4g	-4.0861826	9.81E-14	1.95E-11
MGP_FVBNJ_G0032037	Abcc6	-4.0698209	7.44E-14	1.56E-11
MGP_FVBNJ_G0016754	Slc30a10	-4.0635858	2.83E-11	3.59E-09
MGP_FVBNJ_G0029772	Sds	-4.0512018	3.54E-10	3.41E-08
MGP_FVBNJ_G0021759	BC025446	-4.0497874	2.72E-13	5.05E-11
MGP_FVBNJ_G0027157	Cd5l	-4.0495794	1.23E-11	1.66E-09
MGP_FVBNJ_G0017451	Pah	-4.0485615	2.87E-12	4.42E-10
MGP_FVBNJ_G0024904	Cyp2c38	-4.0449477	8.32E-12	1.17E-09
MGP_FVBNJ_G0004057	Gm42604	-4.034288	1.68E-07	9.43E-06
MGP_FVBNJ_G0017484	Amdhd1	-4.0294955	1.61E-11	2.11E-09
MGP_FVBNJ_G0034503	Apoa1	-4.0234294	2.05E-12	3.26E-10
MGP_FVBNJ_G0019643	NA	-4.015506	2.83E-08	1.82E-06
MGP_FVBNJ_G0033704	NA	-4.0105057	4.55E-14	1.07E-11
MGP_FVBNJ_G0017864	Igfbp1	-3.999106	6.94E-08	4.21E-06
MGP_FVBNJ_G0026347	Fermt1	-3.9990361	1.47E-10	1.54E-08

MGP_FVBNJ_G0002767	Prox1os	-3.9980276	2.06E-06	9.29E-05
MGP_FVBNJ_G0019019	G6pc	-3.9953669	3.98E-09	2.98E-07
MGP_FVBNJ_G0031562	Apoc4	-3.9938355	1.48E-11	1.96E-09
MGP_FVBNJ_G0016809	Iyd	-3.9932927	8.30E-11	9.35E-09
MGP_FVBNJ_G0024465	2010003K11Rik	-3.9773116	5.10E-12	7.68E-10
MGP_FVBNJ_G0021322	Adamdec1	-3.9676911	3.65E-09	2.76E-07
MGP_FVBNJ_G0031563	Apoc1	-3.9540305	7.09E-12	1.01E-09
MGP_FVBNJ_G0020980	Fam25c	-3.9405985	1.32E-10	1.41E-08
MGP_FVBNJ_G0033866	Hp	-3.9392352	2.71E-15	9.44E-13
MGP_FVBNJ_G0034958	1300017J02Rik	-3.9312135	8.13E-12	1.15E-09
MGP_FVBNJ_G0022002	Ppara	-3.9270184	1.10E-11	1.51E-09
MGP_FVBNJ_G0030158	Pdx1	-3.9116946	3.27E-07	1.72E-05
MGP_FVBNJ_G0020303	Adtrp	-3.9077582	1.15E-13	2.27E-11
MGP_FVBNJ_G0021534	Agxt2	-3.8977547	3.76E-11	4.60E-09
MGP_FVBNJ_G0021954	Cyp2d11	-3.8900518	1.33E-10	1.41E-08
MGP_FVBNJ_G0020019	Akr1c6	-3.8899634	9.49E-14	1.92E-11
MGP_FVBNJ_G0019331	Gcgr	-3.8849517	4.93E-11	5.93E-09
MGP_FVBNJ_G0025080	Gfra1	-3.8846429	4.60E-15	1.49E-12
MGP_FVBNJ_G0027119	Tdo2	-3.8789756	2.82E-09	2.20E-07
MGP_FVBNJ_G0018405	Asgr1	-3.8730799	1.80E-10	1.84E-08
MGP_FVBNJ_G0017843	Gck	-3.8652862	1.43E-11	1.91E-09
MGP_FVBNJ_G0035346	Rgn	-3.8583768	2.67E-12	4.16E-10
MGP_FVBNJ_G0020634	NA	-3.8552538	8.44E-09	5.96E-07
MGP_FVBNJ_G0027409	Hfe2	-3.8488237	1.63E-09	1.35E-07
MGP_FVBNJ_G0016118	Ugt1a9	-3.8444069	4.69E-11	5.68E-09
MGP_FVBNJ_G0024909	NA	-3.8415931	6.61E-11	7.60E-09
MGP_FVBNJ_G0018579	Pipox	-3.8397945	1.09E-11	1.51E-09
MGP_FVBNJ_G0031431	Slc27a5	-3.8315239	2.17E-11	2.82E-09
MGP_FVBNJ_G0029577	Hsd17b13	-3.8295717	3.30E-09	2.53E-07
MGP_FVBNJ_G0016746	Mar-01	-3.8223125	1.94E-10	1.96E-08
MGP_FVBNJ_G0019374	Zfp750	-3.816966	2.69E-11	3.43E-09
MGP_FVBNJ_G0001201	NA	-3.8134218	8.78E-07	4.21E-05
MGP_FVBNJ_G0036830	Gm44167	-3.8080727	1.04E-05	0.00039372
MGP_FVBNJ_G0028905	Masp2	-3.8030792	2.77E-10	2.71E-08
MGP_FVBNJ_G0020636	Bhmt2	-3.798478	9.79E-10	8.46E-08
MGP_FVBNJ_G0024912	Cyp2c50	-3.7921773	9.41E-11	1.04E-08
MGP_FVBNJ_G0027181	Rhbg	-3.7865628	6.29E-08	3.83E-06
MGP_FVBNJ_G0027867	Ttpa	-3.7834387	3.31E-10	3.22E-08
MGP_FVBNJ_G0024210	Cdo1	-3.7821915	1.65E-10	1.71E-08
MGP_FVBNJ_G0021214	Gjb2	-3.7595639	4.36E-10	4.09E-08
MGP_FVBNJ_G0005179	1810063I02Rik	-3.7504759	1.09E-05	0.00041064

MGP_FVBNJ_G0028064	Baat	-3.7485198	3.77E-10	3.62E-08
MGP_FVBNJ_G0017217	Ftcd	-3.7442946	7.68E-10	6.74E-08
MGP_FVBNJ_G0018030	Timd2	-3.7422597	3.46E-11	4.28E-09
MGP_FVBNJ_G0029197	Gckr	-3.7336057	5.56E-10	5.04E-08
MGP_FVBNJ_G0030160	Urad	-3.7297987	5.20E-09	3.79E-07
MGP_FVBNJ_G0031660	Cyp2b9	-3.7291994	1.23E-08	8.40E-07
MGP_FVBNJ_G0022610	Hgd	-3.7197365	2.50E-11	3.22E-09
MGP_FVBNJ_G0019711	Slc10a1	-3.7141564	1.17E-10	1.28E-08
MGP_FVBNJ_G0018010	Adra1b	-3.7140842	7.91E-11	8.95E-09
MGP_FVBNJ_G0016589	Apoa2	-3.7090294	8.96E-12	1.24E-09
MGP_FVBNJ_G0001819	NA	-3.7063006	4.10E-05	0.0013074
MGP_FVBNJ_G0023625	Gnmt	-3.6971777	1.88E-07	1.04E-05
MGP_FVBNJ_G0029454	NA	-3.6857988	1.78E-10	1.83E-08
MGP_FVBNJ_G0026909	Slc2a2	-3.6849137	8.26E-10	7.19E-08
MGP_FVBNJ_G0016640	Olfr16	-3.6836052	9.72E-05	0.00269577
MGP_FVBNJ_G0023973	Dsg1c	-3.6835835	2.40E-06	0.00010499
MGP_FVBNJ_G0030141	NA	-3.6798598	1.32E-11	1.77E-09
MGP_FVBNJ_G0003191	Gm16573	-3.6737977	2.49E-06	0.00010744
MGP_FVBNJ_G0019212	Slc16a5	-3.6735286	4.86E-08	3.01E-06
MGP_FVBNJ_G0021854	Tmprss6	-3.6576976	4.54E-10	4.22E-08
MGP_FVBNJ_G0005015	Gm2061	-3.6525562	2.60E-05	0.00088892
MGP_FVBNJ_G0021300	Gulo	-3.6518446	1.94E-09	1.58E-07
MGP_FVBNJ_G0033933	Hsd17b2	-3.6492048	1.39E-09	1.17E-07
MGP_FVBNJ_G0035403	Rhox5	-3.6481228	8.73E-07	4.19E-05
MGP_FVBNJ_G0032963	Cyp2e1	-3.644909	2.56E-10	2.54E-08
MGP_FVBNJ_G0009927	NA	-3.6442557	0.00016225	0.00420116
MGP_FVBNJ_G0028246	Cyp2j5	-3.6433449	2.07E-12	3.27E-10
MGP_FVBNJ_G0024903	Cyp2c29	-3.638457	3.05E-11	3.84E-09
MGP_FVBNJ_G0024962	Abcc2	-3.6278217	3.05E-09	2.37E-07
MGP_FVBNJ_G0033390	Aadat	-3.6240494	8.59E-09	6.04E-07
MGP_FVBNJ_G0019878	Serpina12	-3.6078861	3.78E-09	2.86E-07
MGP_FVBNJ_G0016050	Scg2	-3.6065775	0.00010331	0.00283956
MGP_FVBNJ_G0020200	NA	-3.6044292	1.51E-10	1.58E-08
MGP_FVBNJ_G0020407	F12	-3.6024265	6.16E-09	4.40E-07
MGP_FVBNJ_G0030626	Fabp1	-3.6022425	3.87E-10	3.70E-08
MGP_FVBNJ_G0016933	Smlr1	-3.6017333	4.13E-10	3.94E-08
MGP_FVBNJ_G0016130	Spp2	-3.5916937	4.80E-09	3.53E-07
MGP_FVBNJ_G0017802	Sec14l4	-3.5863095	1.20E-10	1.29E-08
MGP_FVBNJ_G0029461	Sult1e1	-3.5856482	3.33E-05	0.00110011
MGP_FVBNJ_G0024911	Cyp2c54	-3.5837648	1.60E-07	9.05E-06
MGP_FVBNJ_G0029456	Ugt2a3	-3.5818213	4.21E-10	3.99E-08

MGP_FVBNJ_G0028031	Stra6l	-3.576419	9.15E-14	1.87E-11
MGP_FVBNJ_G0024260	Iigp1	-3.573879	7.93E-13	1.38E-10
MGP_FVBNJ_G0035963	Cnksr2	-3.5712498	1.44E-08	9.76E-07
MGP_FVBNJ_G0031659	Cyp2b13	-3.5684891	4.67E-08	2.90E-06
MGP_FVBNJ_G0021996	Smc1b	-3.5623747	3.14E-05	0.00104389
MGP_FVBNJ_G0020908	Il17rb	-3.5613787	1.61E-10	1.68E-08
MGP_FVBNJ_G0023548	Crisp1	-3.548272	0.00020528	0.00513558
MGP_FVBNJ_G0029493	Alb	-3.5481249	3.09E-09	2.39E-07
MGP_FVBNJ_G0015160	Gm20319	-3.5472644	6.05E-08	3.69E-06
MGP_FVBNJ_G0022530	Bdh1	-3.5449726	1.95E-10	1.97E-08
MGP_FVBNJ_G0016768	Prox1	-3.5408393	5.09E-14	1.17E-11
MGP_FVBNJ_G0026043	Bbox1	-3.5396543	1.39E-09	1.17E-07
MGP_FVBNJ_G0002341	NA	-3.5385673	1.76E-06	8.04E-05
MGP_FVBNJ_G0030247	Cftr	-3.536022	1.43E-07	8.12E-06
MGP_FVBNJ_G0021501	C6	-3.5347314	2.22E-16	1.08E-13
MGP_FVBNJ_G0007043	1700049E15Rik	-3.5330123	3.92E-05	0.00126163
MGP_FVBNJ_G0028295	Prkaa2	-3.5228535	6.01E-10	5.40E-08
MGP_FVBNJ_G0021957	NA	-3.5219579	4.45E-06	0.00018331
MGP_FVBNJ_G0023610	Slc22a7	-3.5201504	1.81E-08	1.20E-06
MGP_FVBNJ_G0029326	Slc34a2	-3.5183797	2.16E-10	2.16E-08
MGP_FVBNJ_G0007410	Gm27216	-3.5152865	5.21E-08	3.21E-06
MGP_FVBNJ_G0034250	Tbx20	-3.5150427	5.49E-11	6.47E-09
MGP_FVBNJ_G0028145	Orm1	-3.5088794	2.03E-05	0.00072065
MGP_FVBNJ_G0033782	Ces2e	-3.5077732	1.99E-09	1.61E-07
MGP_FVBNJ_G0031455	Sult2a8	-3.5054655	2.58E-09	2.04E-07
MGP_FVBNJ_G0019887	Serpina3m	-3.4962537	3.46E-09	2.64E-07
MGP_FVBNJ_G0032688	Acsm5	-3.4924519	1.10E-09	9.43E-08
MGP_FVBNJ_G0029350	Klb	-3.4898162	3.80E-09	2.86E-07
MGP_FVBNJ_G0035794	Pcdh11x	-3.486151	0.00042683	0.00950505
MGP_FVBNJ_G0018264	Slc47a1	-3.4780505	1.26E-08	8.54E-07
MGP_FVBNJ_G0028381	NA	-3.4743736	1.28E-05	0.00047391
MGP_FVBNJ_G0021670	Sntb1	-3.4729465	5.55E-12	8.14E-10
MGP_FVBNJ_G0030159	Cdx2	-3.4691779	0.00020495	0.00513558
MGP_FVBNJ_G0028067	Aldob	-3.4659816	1.73E-07	9.67E-06
MGP_FVBNJ_G0025312	Adamts13	-3.4646677	5.22E-08	3.21E-06
MGP_FVBNJ_G0031838	Hpn	-3.4613878	1.91E-09	1.56E-07
MGP_FVBNJ_G0034761	Lipc	-3.4609893	4.46E-10	4.16E-08
MGP_FVBNJ_G0034506	Apoa5	-3.4545073	1.61E-08	1.08E-06
MGP_FVBNJ_G0008081	Gm44482	-3.4490909	4.75E-05	0.00147669
MGP_FVBNJ_G0029055	Steap4	-3.4464319	5.33E-09	3.87E-07
MGP_FVBNJ_G0039850	NA	-3.4431929	3.59E-09	2.73E-07

MGP_FVBNJ_G0015966	Erbb4	-3.4423363	7.12E-07	3.49E-05
MGP_FVBNJ_G0024790	Mamdc2	-3.4400909	2.28E-09	1.82E-07
MGP_FVBNJ_G0030258	Aass	-3.4387733	4.89E-08	3.02E-06
MGP_FVBNJ_G0039683	NA	-3.4286183	7.47E-13	1.31E-10
MGP_FVBNJ_G0025018	Cyp17a1	-3.4247797	0.00014325	0.00374616
MGP_FVBNJ_G0034470	Treh	-3.4236785	3.56E-05	0.0011663
MGP_FVBNJ_G0026630	Rims4	-3.420939	1.44E-09	1.20E-07
MGP_FVBNJ_G0021454	Slitrk6	-3.4157757	0.00061915	0.01298345
MGP_FVBNJ_G0033828	Lcat	-3.4035897	3.61E-11	4.44E-09
MGP_FVBNJ_G0024868	Slc16a12	-3.3948954	1.02E-10	1.13E-08
MGP_FVBNJ_G0027613	Tmem56	-3.3931953	1.02E-09	8.79E-08
MGP_FVBNJ_G0021749	Ly6d	-3.377645	4.69E-09	3.49E-07
MGP_FVBNJ_G0017725	Mettl7b	-3.3758906	3.07E-11	3.84E-09
MGP_FVBNJ_G0028383	NA	-3.375555	4.54E-05	0.00142244
MGP_FVBNJ_G0021041	1700011H14Rik	-3.3713798	1.95E-05	0.0006942
MGP_FVBNJ_G0020909	Chdh	-3.3689446	3.83E-08	2.42E-06
MGP_FVBNJ_G0023752	C3	-3.3685162	2.01E-12	3.22E-10
MGP_FVBNJ_G0028098	NA	-3.3676392	6.87E-10	6.07E-08
MGP_FVBNJ_G0022761	Robo2	-3.3634996	4.78E-09	3.53E-07
MGP_FVBNJ_G0029065	Abcb4	-3.3625707	6.67E-10	5.93E-08
MGP_FVBNJ_G0028318	Dio1	-3.361555	8.46E-09	5.96E-07
MGP_FVBNJ_G0019712	Smoc1	-3.3598391	3.09E-11	3.85E-09
MGP_FVBNJ_G0021958	Cyp2d12	-3.3585809	1.07E-07	6.27E-06
MGP_FVBNJ_G0035027	Gnat1	-3.3529914	4.16E-08	2.59E-06
MGP_FVBNJ_G0026928	Kcnmb2	-3.3484109	9.97E-05	0.00275703
MGP_FVBNJ_G0030987	Slc6a13	-3.3446011	1.83E-08	1.21E-06
MGP_FVBNJ_G0029451	Ugt2b1	-3.3418441	2.08E-09	1.68E-07
MGP_FVBNJ_G0023291	Cbs	-3.3394451	5.53E-12	8.14E-10
MGP_FVBNJ_G0022701	Adgrg7	-3.3359906	4.30E-06	0.00017903
MGP_FVBNJ_G0017301	Elane	-3.3353617	6.75E-05	0.0019871
MGP_FVBNJ_G0022149	Gpd1	-3.3251233	4.78E-09	3.53E-07
MGP_FVBNJ_G0015693	Xkr9	-3.3222088	4.01E-08	2.52E-06
MGP_FVBNJ_G0020637	Dmgdh	-3.3182585	3.03E-08	1.93E-06
MGP_FVBNJ_G0021960	Cyp2d37-ps	-3.3157751	9.78E-08	5.78E-06
MGP_FVBNJ_G0015965	Cps1	-3.3153113	5.80E-09	4.15E-07
MGP_FVBNJ_G0026237	Slc27a2	-3.3075621	7.25E-08	4.38E-06
MGP_FVBNJ_G0024717	Glyat	-3.3044358	1.89E-10	1.92E-08
MGP_FVBNJ_G0022617	Nr1i2	-3.3040188	7.79E-07	3.80E-05
MGP_FVBNJ_G0025258	C8g	-3.3021597	7.28E-10	6.42E-08
MGP_FVBNJ_G0032687	Pdilt	-3.2971918	1.37E-05	0.00050803
MGP_FVBNJ_G0032701	Anks4b	-3.2953208	4.94E-09	3.62E-07

MGP_FVBNJ_G0030257	Ptprz1	-3.2940515	3.16E-06	0.00013397
MGP_FVBNJ_G0016348	Adora1	-3.2936297	6.47E-09	4.60E-07
MGP_FVBNJ_G0029402	Lrrc66	-3.2919097	6.64E-07	3.27E-05
MGP_FVBNJ_G0021961	Cyp2d40	-3.2910767	1.02E-08	7.08E-07
MGP_FVBNJ_G0023314	Cyp4f14	-3.290659	1.05E-09	9.04E-08
MGP_FVBNJ_G0033779	Ces2a	-3.2868973	9.92E-09	6.91E-07
MGP_FVBNJ_G0034630	Cyp1a2	-3.2867336	2.60E-09	2.05E-07
MGP_FVBNJ_G0002859	Gm11967	-3.286694	0.00010192	0.00281036
MGP_FVBNJ_G0034981	Glyctk	-3.2828341	1.39E-10	1.47E-08
MGP_FVBNJ_G0023860	Tmem178	-3.2825317	5.61E-13	1.00E-10
MGP_FVBNJ_G0013269	NA	-3.2800689	9.33E-05	0.00260433
MGP_FVBNJ_G0012262	NA	-3.2751473	2.31E-06	0.00010143
MGP_FVBNJ_G0031115	Clec1b	-3.2708348	7.90E-08	4.72E-06
MGP_FVBNJ_G0021573	Dnah5	-3.2684439	8.80E-05	0.00248535
MGP_FVBNJ_G0030675	Eva1a	-3.2679977	1.06E-10	1.17E-08
MGP_FVBNJ_G0029215	Slc5a1	-3.2658913	2.43E-07	1.32E-05
MGP_FVBNJ_G0019654	Tmem30b	-3.2642712	3.91E-08	2.46E-06
MGP_FVBNJ_G0031229	Gm6614	-3.2640094	3.93E-07	2.03E-05
MGP_FVBNJ_G0026913	Egfem1	-3.2602357	3.80E-06	0.00015883
MGP_FVBNJ_G0031195	Gucy2c	-3.2586621	5.98E-07	2.99E-05
MGP_FVBNJ_G0033874	Tat	-3.2558103	1.68E-07	9.43E-06
MGP_FVBNJ_G0019044	Nags	-3.2514617	2.29E-08	1.49E-06
MGP_FVBNJ_G0027425	Pdzk1	-3.2510466	1.08E-08	7.44E-07
MGP_FVBNJ_G0019921	Slc25a47	-3.2496782	3.46E-07	1.82E-05
MGP_FVBNJ_G0023282	Tff1	-3.2463616	0.00016252	0.00420116
MGP_FVBNJ_G0027297	Ivl	-3.2419803	4.51E-06	0.00018525
MGP_FVBNJ_G0033911	Cntnap4	-3.2287388	0.00035986	0.00819377
MGP_FVBNJ_G0007400	Gm28182	-3.2242943	5.75E-05	0.00174066
MGP_FVBNJ_G0027904	Cnr1	-3.2203041	0.0005504	0.01175187
MGP_FVBNJ_G0020015	Akr1c18	-3.217043	0.00074564	0.01513477
MGP_FVBNJ_G0021237	Shisa2	-3.2154143	4.00E-11	4.87E-09
MGP_FVBNJ_G0027946	Dnaic1	-3.2146207	6.37E-06	0.00025419
MGP_FVBNJ_G0030140	NA	-3.2099454	1.22E-10	1.31E-08
MGP_FVBNJ_G0031237	Gys2	-3.20643	1.35E-06	6.27E-05
MGP_FVBNJ_G0015441	5033403H07Rik	-3.202546	5.99E-07	2.99E-05
MGP_FVBNJ_G0039857	NA	-3.2011742	1.38E-07	7.92E-06
MGP_FVBNJ_G0023547	Crisp3	-3.1960234	0.0012305	0.02263124
MGP_FVBNJ_G0034787	Onecut1	-3.1839161	1.80E-07	1.01E-05
MGP_FVBNJ_G0021660	Mal2	-3.1791008	1.21E-09	1.02E-07
MGP_FVBNJ_G0027816	Cyp7a1	-3.1787186	2.23E-08	1.45E-06
MGP_FVBNJ_G0024466	Pcx	-3.1749486	1.88E-09	1.55E-07

MGP_FVBNJ_G0024908	NA	-3.1711447	1.71E-10	1.76E-08
MGP_FVBNJ_G0020398	Fgfr4	-3.1690929	1.63E-15	6.14E-13
MGP_FVBNJ_G0023709	Lrg1	-3.1671682	5.44E-11	6.45E-09
MGP_FVBNJ_G0029099	Reln	-3.163207	2.20E-09	1.76E-07
MGP_FVBNJ_G0025520	Kynu	-3.1614426	1.59E-11	2.10E-09
MGP_FVBNJ_G0028289	Dab1	-3.1597648	6.27E-06	0.00025038
MGP_FVBNJ_G0004158	2210019I11Rik	-3.1596637	8.79E-05	0.00248535
MGP_FVBNJ_G0026185	Ckmt1	-3.151456	2.67E-05	0.00090976
MGP_FVBNJ_G0024837	Mbl2	-3.1489399	5.49E-09	3.96E-07
MGP_FVBNJ_G0003631	Hnf4aos	-3.1485209	0.00010814	0.00295407
MGP_FVBNJ_G0016306	Pm20d1	-3.1468612	6.47E-10	5.79E-08
MGP_FVBNJ_G0027710	Adh1	-3.1403217	2.69E-08	1.74E-06
MGP_FVBNJ_G0018464	Ggt6	-3.1394984	1.30E-07	7.50E-06
MGP_FVBNJ_G0030592	Gprin3	-3.138731	2.46E-09	1.95E-07
MGP_FVBNJ_G0029460	Sult1d1	-3.1370362	5.25E-07	2.68E-05
MGP_FVBNJ_G0016333	Etnk2	-3.1269988	5.39E-07	2.73E-05
MGP_FVBNJ_G0029448	Tmprss11e	-3.122105	0.00204746	0.03360823
MGP_FVBNJ_G0000128	NA	-3.1215958	0.00072314	0.01477504
MGP_FVBNJ_G0025996	Slc1a2	-3.120714	7.52E-08	4.52E-06
MGP_FVBNJ_G0032693	Acsm3	-3.1095904	8.64E-08	5.14E-06
MGP_FVBNJ_G0033382	Hand2	-3.1072173	5.87E-15	1.75E-12
MGP_FVBNJ_G0030708	Cd207	-3.1067452	1.72E-05	0.00062094
MGP_FVBNJ_G0028166	Rasef	-3.0956926	3.60E-06	0.00015166
MGP_FVBNJ_G0022747	Epha3	-3.0856953	1.10E-10	1.20E-08
MGP_FVBNJ_G0034505	Apoa4	-3.0849861	0.000102	0.00281036
MGP_FVBNJ_G0021319	Nefm	-3.0794147	0.00052972	0.01141962
MGP_FVBNJ_G0030485	Aoc1	-3.0761429	6.65E-05	0.00195978
MGP_FVBNJ_G0027627	Fabp2	-3.0753203	5.13E-06	0.00020923
MGP_FVBNJ_G0034801	Bmp5	-3.0708717	1.30E-10	1.40E-08
MGP_FVBNJ_G0025412	Lcn2	-3.0671256	1.25E-07	7.25E-06
MGP_FVBNJ_G0018690	Gm11437	-3.0664676	1.07E-05	0.00040569
MGP_FVBNJ_G0031953	NA	-3.0648541	6.52E-10	5.82E-08
MGP_FVBNJ_G0016405	Kcnt2	-3.0622519	7.42E-11	8.48E-09
MGP_FVBNJ_G0025455	AI182371	-3.0609766	1.17E-08	7.97E-07
MGP_FVBNJ_G0016465	Fam163a	-3.0598062	6.03E-05	0.0018086
MGP_FVBNJ_G0016905	Tcf21	-3.058837	9.12E-11	1.02E-08
MGP_FVBNJ_G0031217	Pik3c2g	-3.0571101	1.07E-07	6.27E-06
MGP_FVBNJ_G0019879	Serpina5	-3.0444655	0.00067332	0.0139101
MGP_FVBNJ_G0025068	Plekhs1	-3.0437019	1.95E-06	8.81E-05
MGP_FVBNJ_G0020229	Agtr1a	-3.0427372	1.54E-09	1.28E-07
MGP_FVBNJ_G0030764	Bmp10	-3.0416389	2.54E-06	0.00010914

MGP_FVBNJ_G0029077	Hgf	-3.0379811	8.75E-13	1.50E-10
MGP_FVBNJ_G0030496	Npy	-3.0344718	9.12E-05	0.00256084
MGP_FVBNJ_G0019871	Serpina6	-3.0286357	1.50E-08	1.01E-06
MGP_FVBNJ_G0021529	Capsl	-3.0220367	3.91E-05	0.00126163
MGP_FVBNJ_G0020187	Slc17a4	-3.0105703	2.40E-09	1.91E-07
MGP_FVBNJ_G0016391	Nr5a2	-3.0063291	9.69E-09	6.77E-07
MGP_FVBNJ_G0018761	Hlf	-3.0026993	2.21E-06	9.80E-05
MGP_FVBNJ_G0025214	Il1f10	-3.0024147	7.23E-05	0.00211482
MGP_FVBNJ_G0006518	Gm15611	-2.9946829	3.39E-05	0.00111515
MGP_FVBNJ_G0016909	Vnn3	-2.9944311	6.49E-08	3.94E-06
MGP_FVBNJ_G0030812	Aldh1l1	-2.9910843	2.00E-08	1.31E-06
MGP_FVBNJ_G0040130	NA	-2.9897183	1.66E-05	0.00060289
MGP_FVBNJ_G0023708	Plin5	-2.9815376	2.17E-05	0.00076201
MGP_FVBNJ_G0025145	Itih2	-2.979824	4.02E-09	3.00E-07
MGP_FVBNJ_G0003818	Gm11844	-2.9748435	1.54E-05	0.00056353
MGP_FVBNJ_G0017548	Syt1	-2.9685377	2.43E-06	0.00010598
MGP_FVBNJ_G0019163	Abca8b	-2.9670726	3.14E-09	2.41E-07
MGP_FVBNJ_G0027726	Gbp2b	-2.9634567	3.65E-07	1.90E-05
MGP_FVBNJ_G0003653	Ctcflos	-2.9572027	1.43E-05	0.00052611
MGP_FVBNJ_G0034081	Mmp10	-2.9531675	0.00159665	0.02769713
MGP_FVBNJ_G0020968	Gdf2	-2.94748	3.01E-06	0.00012847
MGP_FVBNJ_G0020616	Acot12	-2.9467399	1.25E-05	0.0004662
MGP_FVBNJ_G0004152	Gm16035	-2.9457635	0.00014335	0.00374616
MGP_FVBNJ_G0017370	Creb3l3	-2.9424127	1.67E-07	9.40E-06
MGP_FVBNJ_G0033234	Ido2	-2.9393847	4.17E-08	2.59E-06
MGP_FVBNJ_G0023971	Dsc2	-2.9391124	8.81E-09	6.17E-07
MGP_FVBNJ_G0031271	Pthlh	-2.9362211	0.00063317	0.01319066
MGP_FVBNJ_G0001366	NA	-2.9362103	6.50E-07	3.21E-05
MGP_FVBNJ_G0027447	Hsd3b1	-2.9359246	8.60E-06	0.00033297
MGP_FVBNJ_G0005720	1700001G01Rik	-2.9349011	6.89E-05	0.00202285
MGP_FVBNJ_G0035202	Vipr1	-2.9327118	2.49E-07	1.35E-05
MGP_FVBNJ_G0027961	NA	-2.926118	8.60E-06	0.00033297
MGP_FVBNJ_G0018808	Ngfr	-2.91696	3.48E-06	0.00014678
MGP_FVBNJ_G0032350	Plekhb1	-2.9109646	8.39E-08	5.00E-06
MGP_FVBNJ_G0021977	Scube1	-2.9103017	2.14E-06	9.58E-05
MGP_FVBNJ_G0011575	NA	-2.9077087	1.50E-05	0.00055085
MGP_FVBNJ_G0023502	Ubd	-2.903389	8.91E-05	0.00250986
MGP_FVBNJ_G0032569	Syt9	-2.9008709	2.91E-07	1.54E-05
MGP_FVBNJ_G0016546	Gpa33	-2.8995602	0.00312518	0.04628793
MGP_FVBNJ_G0006798	Gm27033	-2.8986822	0.00080492	0.0160631
MGP_FVBNJ_G0020791	Dnase1l3	-2.8940036	1.88E-08	1.24E-06

MGP_FVBNJ_G0033706	Ces1e	-2.8920356	0.00034561	0.00790157
MGP_FVBNJ_G0006420	0610043K17Rik	-2.8908217	2.33E-05	0.00080754
MGP_FVBNJ_G0015817	Slc9a2	-2.8881227	0.00012727	0.00339326
MGP_FVBNJ_G0010089	NA	-2.8833464	0.00058868	0.01241459
MGP_FVBNJ_G0012887	NA	-2.8803248	0.00040912	0.00917473
MGP_FVBNJ_G0017122	NA	-2.8779934	2.75E-07	1.46E-05
MGP_FVBNJ_G0028126	NA	-2.870418	5.01E-10	4.60E-08
MGP_FVBNJ_G0021193	Mcpt2	-2.8698979	0.00295647	0.04440881
MGP_FVBNJ_G0030388	Mgam	-2.8695525	1.57E-07	8.89E-06
MGP_FVBNJ_G0017383	Tjp3	-2.867365	7.80E-10	6.82E-08
MGP_FVBNJ_G0022896	Dscam	-2.8657423	0.00020929	0.00520686
MGP_FVBNJ_G0024910	Cyp2c37	-2.8389997	5.75E-07	2.90E-05
MGP_FVBNJ_G0028382	NA	-2.834547	0.00055537	0.01183537
MGP_FVBNJ_G0040280	NA	-2.8321092	0.00010668	0.00292125
MGP_FVBNJ_G0022038	Tymp	-2.823468	1.25E-07	7.28E-06
MGP_FVBNJ_G0020694	Ocln	-2.8214638	3.09E-06	0.00013132
MGP_FVBNJ_G0016248	Sctr	-2.8199973	0.000113	0.00306058
MGP_FVBNJ_G0016992	Mfsd4b1	-2.8164847	4.32E-06	0.00017951
MGP_FVBNJ_G0003846	NA	-2.8144907	8.88E-06	0.00034144
MGP_FVBNJ_G0034205	Ap1m2	-2.8131716	5.62E-06	0.00022639
MGP_FVBNJ_G0031317	Eps8l1	-2.8098304	2.12E-05	0.00074896
MGP_FVBNJ_G0017977	Gabrp	-2.8015557	0.00010489	0.00287586
MGP_FVBNJ_G0027838	Cdh17	-2.8014058	1.22E-05	0.00045618
MGP_FVBNJ_G0015796	Aff3	-2.7987984	6.43E-07	3.18E-05
MGP_FVBNJ_G0032985	Ano9	-2.7969105	2.20E-05	0.00076816
MGP_FVBNJ_G0018118	Leap2	-2.796882	4.41E-05	0.00138604
MGP_FVBNJ_G0034957	Trf	-2.7928708	2.56E-10	2.54E-08
MGP_FVBNJ_G0022967	Dact2	-2.7917107	5.55E-11	6.51E-09
MGP_FVBNJ_G0016999	Ddo	-2.7904822	4.84E-07	2.48E-05
MGP_FVBNJ_G0020806	Cadps	-2.7885177	0.00219353	0.03531745
MGP_FVBNJ_G0019641	Lrrc9	-2.7865592	0.00035558	0.00811276
MGP_FVBNJ_G0022471	Ehhadh	-2.7865032	2.53E-05	0.00086756
MGP_FVBNJ_G0032050	Saa2	-2.7826569	3.12E-09	2.41E-07
MGP_FVBNJ_G0028361	Slc5a9	-2.7816143	0.00135511	0.02432106
MGP_FVBNJ_G0028157	Astn2	-2.7806686	4.05E-07	2.09E-05
MGP_FVBNJ_G0010363	NA	-2.779573	0.00273932	0.04179475
MGP_FVBNJ_G0032075	Csrp3	-2.7752775	4.40E-06	0.00018196
MGP_FVBNJ_G0019427	Mycn	-2.7643635	5.76E-11	6.69E-09
MGP_FVBNJ_G0016560	Rxrg	-2.7617904	5.56E-05	0.00169175
MGP_FVBNJ_G0015990	Gpbar1	-2.7579464	1.12E-05	0.00042091
MGP_FVBNJ_G0031568	Cblc	-2.754922	7.44E-07	3.64E-05

MGP_FVBNJ_G0022794	Cldn8	-2.7511226	0.00020277	0.00509586
MGP_FVBNJ_G0033785	Ces2h	-2.7511061	0.00033957	0.00779545
MGP_FVBNJ_G0032663	Sox6	-2.7482308	5.48E-09	3.96E-07
MGP_FVBNJ_G0025393	NA	-2.7475569	7.83E-07	3.81E-05
MGP_FVBNJ_G0019960	Aspg	-2.7472239	2.23E-06	9.88E-05
MGP_FVBNJ_G0025603	Abcb11	-2.7454287	8.44E-07	4.08E-05
MGP_FVBNJ_G0019466	Colec11	-2.7270294	9.29E-07	4.44E-05
MGP_FVBNJ_G0007249	2810030D12Rik	-2.7236966	0.00015768	0.00410121
MGP_FVBNJ_G0018258	Kcnj12	-2.721736	0.00249756	0.03917878
MGP_FVBNJ_G0017805	Sec14l2	-2.7171784	2.15E-07	1.18E-05
MGP_FVBNJ_G0021532	NA	-2.7164939	5.84E-07	2.93E-05
MGP_FVBNJ_G0025156	Itga8	-2.7146656	1.45E-06	6.71E-05
MGP_FVBNJ_G0031332	Il11	-2.7135463	0.00096976	0.01865142
MGP_FVBNJ_G0034047	BC021891	-2.7094717	2.29E-07	1.26E-05
MGP_FVBNJ_G0027456	Vtcn1	-2.7064708	2.23E-06	9.87E-05
MGP_FVBNJ_G0021181	Cideb	-2.7063088	1.38E-07	7.92E-06
MGP_FVBNJ_G0016963	Rsph4a	-2.7023866	6.42E-06	0.00025554
MGP_FVBNJ_G0030911	Atp2b2	-2.7018701	1.72E-05	0.00062116
MGP_FVBNJ_G0018967	Krt14	-2.6963147	0.00121858	0.02244901
MGP_FVBNJ_G0033045	Slc22a18	-2.6904751	4.02E-08	2.52E-06
MGP_FVBNJ_G0002815	Gm15343	-2.6897998	4.61E-05	0.00143903
MGP_FVBNJ_G0028245	Cyp2j9	-2.6891395	7.89E-05	0.00227343
MGP_FVBNJ_G0028294	1700024P16Rik	-2.6867132	9.87E-06	0.00037612
MGP_FVBNJ_G0034713	Slc51b	-2.6822833	0.00055973	0.01188274
MGP_FVBNJ_G0026356	Snap25	-2.680576	4.88E-05	0.00151466
MGP_FVBNJ_G0024655	Vwce	-2.6751615	2.82E-05	0.00094812
MGP_FVBNJ_G0022008	Celsr1	-2.6748818	1.10E-18	1.23E-15
MGP_FVBNJ_G0032685	Gp2	-2.673799	0.00094263	0.01822384
MGP_FVBNJ_G0029887	Glt1d1	-2.6721405	0.00057077	0.01208263
MGP_FVBNJ_G0030988	Slc6a12	-2.6698834	2.48E-06	0.0001074
MGP_FVBNJ_G0024980	Pax2	-2.6682266	0.00044803	0.0099175
MGP_FVBNJ_G0023681	Kcnh8	-2.6647051	0.00188221	0.03147505
MGP_FVBNJ_G0005690	Gm19585	-2.6630068	0.00038266	0.00865987
MGP_FVBNJ_G0033708	Ces1g	-2.6576059	5.86E-07	2.94E-05
MGP_FVBNJ_G0019995	Ptprn2	-2.6551499	0.0001027	0.00282638
MGP_FVBNJ_G0016402	NA	-2.6541398	1.96E-06	8.83E-05
MGP_FVBNJ_G0030246	Asz1	-2.653149	0.00095377	0.01842322
MGP_FVBNJ_G0003327	Gm15318	-2.652406	1.62E-05	0.00059021
MGP_FVBNJ_G0031674	Cyp2f2	-2.648899	1.54E-08	1.03E-06
MGP_FVBNJ_G0030530	Wipf3	-2.6452988	1.31E-08	8.90E-07
MGP_FVBNJ_G0003059	Gm11789	-2.636573	3.39E-05	0.00111548

MGP_FVBNJ_G0023662	Treml4	-2.6349311	1.24E-05	0.00046369
MGP_FVBNJ_G0032211	Fsd2	-2.626727	0.0003315	0.00761796
MGP_FVBNJ_G0027584	NA	-2.6258563	2.66E-05	0.00090796
MGP_FVBNJ_G0027441	Hmgcs2	-2.6247123	2.39E-06	0.00010465
MGP_FVBNJ_G0016675	Kmo	-2.6201989	7.40E-08	4.46E-06
MGP_FVBNJ_G0028450	2610528J11Rik	-2.6195129	9.68E-05	0.00268686
MGP_FVBNJ_G0003011	Gm11586	-2.6191837	0.00023381	0.00570861
MGP_FVBNJ_G0030202	NA	-2.6010732	3.17E-05	0.00105056
MGP_FVBNJ_G0027464	Mab21l3	-2.6004439	4.69E-06	0.00019206
MGP_FVBNJ_G0033932	Sdr42e1	-2.5988944	4.02E-06	0.0001677
MGP_FVBNJ_G0023869	Haao	-2.5944823	1.07E-06	5.11E-05
MGP_FVBNJ_G0005577	Gm28876	-2.5932614	0.00148733	0.02629196
MGP_FVBNJ_G0019868	Ifi27l2b	-2.5894107	1.16E-06	5.52E-05
MGP_FVBNJ_G0027963	Gm2564	-2.5870921	5.79E-05	0.00174534
MGP_FVBNJ_G0018645	Ccl11	-2.5855537	0.00082808	0.01645161
MGP_FVBNJ_G0018917	Krt23	-2.5842886	4.73E-05	0.00147224
MGP_FVBNJ_G0017245	Lrrc3	-2.5826501	2.55E-07	1.37E-05
MGP_FVBNJ_G0024202	Kcnn2	-2.5752267	1.72E-06	7.86E-05
MGP_FVBNJ_G0021007	Sftpd	-2.5702612	0.00054331	0.01163401
MGP_FVBNJ_G0019186	Gprc5c	-2.5658603	5.33E-10	4.85E-08
MGP_FVBNJ_G0031837	Fxyd3	-2.5644747	1.14E-05	0.00042689
MGP_FVBNJ_G0007272	NA	-2.5635734	2.22E-08	1.45E-06
MGP_FVBNJ_G0031126	Klrc1	-2.562014	1.29E-05	0.00047705
MGP_FVBNJ_G0028961	Ajap1	-2.5575979	0.00110213	0.02069715
MGP_FVBNJ_G0028786	Rnf186	-2.5572878	0.0007252	0.01479343
MGP_FVBNJ_G0031451	Sult2a3	-2.5545292	0.00083026	0.01648019
MGP_FVBNJ_G0024435	NA	-2.5540964	0.0014997	0.02642665
MGP_FVBNJ_G0021197	Ctsg	-2.5506962	0.00264433	0.04082047
MGP_FVBNJ_G0018044	NA	-2.5505685	2.27E-05	0.00079113
MGP_FVBNJ_G0008149	Gm27463	-2.5503567	0.0001306	0.00346166
MGP_FVBNJ_G0032321	Dgat2	-2.5481497	2.14E-06	9.58E-05
MGP_FVBNJ_G0011350	NA	-2.5452556	2.46E-06	0.00010663
MGP_FVBNJ_G0017927	5730522E02Rik	-2.5407601	4.47E-05	0.00140221
MGP_FVBNJ_G0022647	Cfap44	-2.5405847	0.00223743	0.03595856
MGP_FVBNJ_G0027531	Prok1	-2.5392372	0.00020365	0.00511201
MGP_FVBNJ_G0030260	Cadps2	-2.5345071	1.94E-09	1.58E-07
MGP_FVBNJ_G0030349	Akr1d1	-2.5314	9.90E-05	0.00274182
MGP_FVBNJ_G0028507	Mfsd2a	-2.5267975	0.00092593	0.01797906
MGP_FVBNJ_G0023150	Prss29	-2.5264717	0.00109899	0.02065556
MGP_FVBNJ_G0030261	Rnf133	-2.5263621	0.00253889	0.03963191
MGP_FVBNJ_G0001635	NA	-2.5248865	0.00116621	0.02162732

MGP_FVBNJ_G0018734	Mpo	-2.517544	0.00214213	0.03472739
MGP_FVBNJ_G0021613	Grhl2	-2.5172514	3.05E-06	0.00013013
MGP_FVBNJ_G0007037	2010013B24Rik	-2.5171767	0.0009927	0.01905972
MGP_FVBNJ_G0007097	Gm28076	-2.5162152	0.00125987	0.02303837
MGP_FVBNJ_G0003102	Gm15523	-2.5132184	0.00136104	0.02440775
MGP_FVBNJ_G0029591	NA	-2.5082877	1.87E-05	0.00067004
MGP_FVBNJ_G0028591	Hpca	-2.5079766	7.57E-05	0.00218767
MGP_FVBNJ_G0018969	Krt17	-2.5078031	0.00271093	0.04150706
MGP_FVBNJ_G0017670	NA	-2.507134	7.53E-06	0.0002966
MGP_FVBNJ_G0027446	NA	-2.5067092	6.09E-05	0.00182272
MGP_FVBNJ_G0005763	AW112010	-2.5057269	1.43E-07	8.12E-06
MGP_FVBNJ_G0027947	Enho	-2.5027054	5.85E-06	0.00023453
MGP_FVBNJ_G0016671	Fmn2	-2.5021916	0.00042091	0.00939194
MGP_FVBNJ_G0027820	Car8	-2.5019778	7.90E-06	0.00030899
MGP_FVBNJ_G0024258	Gm4951	-2.4971873	2.36E-07	1.29E-05
MGP_FVBNJ_G0021198	Gzme	-2.4949532	0.00336169	0.04891339
MGP_FVBNJ_G0025333	Cel	-2.4918238	0.003016	0.04506014
MGP_FVBNJ_G0019551	Akap6	-2.4892956	2.14E-07	1.18E-05
MGP_FVBNJ_G0023277	Glp1r	-2.4887538	0.00187456	0.0314208
MGP_FVBNJ_G0000718	NA	-2.4862049	8.57E-05	0.00243685
MGP_FVBNJ_G0039557	NA	-2.4758301	5.42E-05	0.00165705
MGP_FVBNJ_G0026349	Hao1	-2.4748041	4.09E-05	0.00130468
MGP_FVBNJ_G0018649	Tmem132e	-2.474634	4.42E-05	0.00138659
MGP_FVBNJ_G0022310	Grin2a	-2.4743102	0.00135268	0.02430771
MGP_FVBNJ_G0022834	Hunk	-2.4732162	6.34E-13	1.12E-10
MGP_FVBNJ_G0039682	NA	-2.4717207	4.36E-10	4.09E-08
MGP_FVBNJ_G0024024	Proc	-2.4699726	2.26E-06	9.97E-05
MGP_FVBNJ_G0001472	NA	-2.4657255	0.00151801	0.02666509
MGP_FVBNJ_G0027122	Gucy1a3	-2.4654909	5.63E-12	8.15E-10
MGP_FVBNJ_G0020967	Gdf10	-2.4621438	4.92E-05	0.00152567
MGP_FVBNJ_G0026259	Prom2	-2.4546962	0.00024106	0.00584748
MGP_FVBNJ_G0039935	NA	-2.4530765	0.0027005	0.04140079
MGP_FVBNJ_G0021538	Ttc23l	-2.452616	0.00248631	0.03910878
MGP_FVBNJ_G0020546	Ube2ql1	-2.4523869	0.00194853	0.03231745
MGP_FVBNJ_G0036719	Gm43189	-2.4511279	0.00024958	0.00602101
MGP_FVBNJ_G0029142	En2	-2.450399	0.00037206	0.00844393
MGP_FVBNJ_G0016635	Slamf8	-2.4491958	0.00025331	0.00605205
MGP_FVBNJ_G0035909	Il13ra2	-2.4491324	0.0020615	0.03376413
MGP_FVBNJ_G0029944	Mlxipl	-2.4483538	1.41E-05	0.00051881
MGP_FVBNJ_G0016404	NA	-2.4466311	2.07E-10	2.07E-08
MGP_FVBNJ_G0029590	NA	-2.4465439	4.24E-05	0.00134129

MGP_FVBNJ_G0016366	Elf3	-2.4462553	1.65E-06	7.60E-05
MGP_FVBNJ_G0030662	Ctnna2	-2.4420923	0.00039863	0.00896659
MGP_FVBNJ_G0018238	Pemt	-2.4282757	1.32E-06	6.16E-05
MGP_FVBNJ_G0027819	Tox	-2.426531	3.45E-06	0.00014585
MGP_FVBNJ_G0026256	Gpat2	-2.4264676	0.00018486	0.00469322
MGP_FVBNJ_G0025507	Lhx2	-2.4248038	3.78E-07	1.96E-05
MGP_FVBNJ_G0004731	Gm5524	-2.4227589	5.02E-06	0.00020513
MGP_FVBNJ_G0002060	NA	-2.4224582	0.00133038	0.02403208
MGP_FVBNJ_G0022386	Serpind1	-2.4218509	5.63E-06	0.00022661
MGP_FVBNJ_G0030262	Rnf148	-2.4182293	0.00173615	0.02958755
MGP_FVBNJ_G0004143	Gm17112	-2.4137283	7.41E-05	0.00215123
MGP_FVBNJ_G0019463	Sox11	-2.4117403	0.00224269	0.03599129
MGP_FVBNJ_G0024720	Keg1	-2.4082599	2.48E-05	0.0008546
MGP_FVBNJ_G0021652	Aard	-2.4057057	5.32E-05	0.00163201
MGP_FVBNJ_G0018877	Ppp1r1b	-2.4050354	0.00065383	0.01358292
MGP_FVBNJ_G0034947	Ky	-2.4049553	0.0012367	0.02270785
MGP_FVBNJ_G0027125	Rbm46	-2.4041529	1.81E-05	0.00065113
MGP_FVBNJ_G0002376	NA	-2.3967965	0.00060005	0.01263049
MGP_FVBNJ_G0035782	Pof1b	-2.3961468	3.06E-05	0.00102105
MGP_FVBNJ_G0022472	Map3k13	-2.3935042	5.34E-07	2.71E-05
MGP_FVBNJ_G0015934	Zdbf2	-2.3927407	0.00345281	0.04997832
MGP_FVBNJ_G0028331	Zyg11a	-2.3923808	0.00011914	0.00319593
MGP_FVBNJ_G0016525	F5	-2.3902138	2.16E-06	9.63E-05
MGP_FVBNJ_G0019430	Fam84a	-2.3894889	2.10E-06	9.45E-05
MGP_FVBNJ_G0018032	Timd4	-2.3885068	7.53E-06	0.0002966
MGP_FVBNJ_G0028046	Gabbr2	-2.3860029	2.43E-05	0.00083779
MGP_FVBNJ_G0020202	Dcdc2a	-2.3787337	0.00017279	0.00442711
MGP_FVBNJ_G0035111	Pth1r	-2.376189	1.21E-06	5.71E-05
MGP_FVBNJ_G0020003	Macc1	-2.3753137	0.00054012	0.01159246
MGP_FVBNJ_G0017128	Hkdc1	-2.374404	0.00022152	0.00545025
MGP_FVBNJ_G0023785	Lama1	-2.3738546	5.34E-07	2.71E-05
MGP_FVBNJ_G0023398	Slc44a4	-2.3703767	7.76E-06	0.0003041
MGP_FVBNJ_G0015725	Gsta3	-2.3699459	9.45E-06	0.00036117
MGP_FVBNJ_G0023123	Meiob	-2.3692215	4.22E-05	0.00133654
MGP_FVBNJ_G0025316	Adamtsl2	-2.3686515	6.22E-05	0.00185035
MGP_FVBNJ_G0017627	Slc16a7	-2.3673906	5.58E-06	0.00022519
MGP_FVBNJ_G0033578	Hhip	-2.3648426	0.00090868	0.01773677
MGP_FVBNJ_G0013643	NA	-2.361933	0.00043874	0.0097333
MGP_FVBNJ_G0025995	Pamr1	-2.3600053	3.90E-08	2.46E-06
MGP_FVBNJ_G0015842	Slc40a1	-2.357568	1.41E-06	6.55E-05
MGP_FVBNJ_G0029338	0610040J01Rik	-2.3573133	1.33E-06	6.20E-05

MGP_FVBNJ_G0002056	NA	-2.3565805	3.27E-07	1.72E-05
MGP_FVBNJ_G0016062	Col4a4	-2.3564771	0.00016361	0.00422109
MGP_FVBNJ_G0017706	Erbb3	-2.3563405	2.34E-07	1.28E-05
MGP_FVBNJ_G0016126	Ugt1a1	-2.3462721	4.03E-05	0.00129387
MGP_FVBNJ_G0027433	Fmo5	-2.3462149	9.92E-06	0.00037725
MGP_FVBNJ_G0028303	Ttc22	-2.3441628	0.0004343	0.00965204
MGP_FVBNJ_G0002888	Gm12195	-2.3416881	0.00199951	0.03291803
MGP_FVBNJ_G0006337	Gm12354	-2.3412233	0.00184452	0.03107951
MGP_FVBNJ_G0025053	Rbm20	-2.3402028	4.12E-05	0.0013125
MGP_FVBNJ_G0027547	Gstm3	-2.3310813	8.30E-05	0.00237462
MGP_FVBNJ_G0004549	NA	-2.3302738	6.19E-05	0.00184465
MGP_FVBNJ_G0034220	Angptl8	-2.3276079	3.53E-05	0.00115833
MGP_FVBNJ_G0029459	Sult1b1	-2.3250465	0.00017319	0.00443236
MGP_FVBNJ_G0030274	Tmem229a	-2.3227021	7.32E-07	3.59E-05
MGP_FVBNJ_G0019475	Sntg2	-2.3208045	0.00013152	0.0034819
MGP_FVBNJ_G0002050	NA	-2.3199037	0.00159927	0.02769955
MGP_FVBNJ_G0021005	Sftpa1	-2.3168524	0.00271724	0.04157172
MGP_FVBNJ_G0021056	Tlr11	-2.3162887	0.0006047	0.01271633
MGP_FVBNJ_G0022317	Ciita	-2.3160876	1.69E-05	0.00061389
MGP_FVBNJ_G0017011	NA	-2.3152021	4.19E-05	0.00133086
MGP_FVBNJ_G0019918	Degs2	-2.3127426	0.00021321	0.0052808
MGP_FVBNJ_G0020186	Slc17a1	-2.3122962	0.00011163	0.00303588
MGP_FVBNJ_G0032966	Urah	-2.311243	9.36E-06	0.00035836
MGP_FVBNJ_G0027121	Gucy1b3	-2.3038416	8.00E-13	1.38E-10
MGP_FVBNJ_G0033562	Nr3c2	-2.2982586	8.36E-07	4.05E-05
MGP_FVBNJ_G0004806	Gm8883	-2.2933997	1.45E-06	6.71E-05
MGP_FVBNJ_G0006035	9430093N23Rik	-2.2902823	0.00327339	0.04794379
MGP_FVBNJ_G0023975	NA	-2.2900588	0.00100735	0.01927457
MGP_FVBNJ_G0035978	Rai2	-2.289746	2.96E-08	1.90E-06
MGP_FVBNJ_G0007463	9530059O14Rik	-2.2876529	0.00327672	0.04795884
MGP_FVBNJ_G0021491	Itgbl1	-2.283698	0.00159792	0.02769763
MGP_FVBNJ_G0031952	Myh14	-2.2804216	0.00037236	0.00844393
MGP_FVBNJ_G0017203	Chchd10	-2.2794085	1.87E-05	0.00066947
MGP_FVBNJ_G0031212	Mgst1	-2.2771033	1.30E-05	0.00048188
MGP_FVBNJ_G0024259	F830016B08Rik	-2.2748686	6.60E-06	0.00026241
MGP_FVBNJ_G0020743	Ankrd55	-2.2706141	2.45E-06	0.00010636
MGP_FVBNJ_G0024443	BC021614	-2.2698358	2.13E-05	0.00074896
MGP_FVBNJ_G0030539	Crhr2	-2.2686167	0.00242079	0.03840706
MGP_FVBNJ_G0034475	Ttc36	-2.2683344	2.74E-05	0.0009274
MGP_FVBNJ_G0030421	Epha1	-2.2676521	3.75E-05	0.00122003
MGP_FVBNJ_G0030195	Tfpi2	-2.2669771	1.47E-06	6.78E-05

MGP_FVBNJ_G0032702	Crym	-2.2630528	0.00019713	0.00497635
MGP_FVBNJ_G0035943	Spin2c	-2.2625927	2.74E-07	1.46E-05
MGP_FVBNJ_G0034913	Clstn2	-2.2609925	5.69E-06	0.00022845
MGP_FVBNJ_G0019892	Clmn	-2.2605605	7.99E-06	0.00031209
MGP_FVBNJ_G0033702	Ces1a	-2.253582	0.00215616	0.0348787
MGP_FVBNJ_G0023365	H2-Ob	-2.2508549	1.54E-05	0.00056352
MGP_FVBNJ_G0021500	Plcxd3	-2.2472755	0.00115581	0.0214701
MGP_FVBNJ_G0033784	Ces2g	-2.2435216	2.94E-07	1.55E-05
MGP_FVBNJ_G0006320	Gm20752	-2.242923	2.41E-06	0.00010511
MGP_FVBNJ_G0002787	Gm29571	-2.2417082	0.00320409	0.04717466
MGP_FVBNJ_G0014965	Gm16565	-2.2402924	0.00016862	0.00434021
MGP_FVBNJ_G0016888	Map7	-2.2390764	7.61E-05	0.00219739
MGP_FVBNJ_G0022053	Kif21a	-2.2378074	2.52E-07	1.36E-05
MGP_FVBNJ_G0028103	D630039A03Rik	-2.2369796	6.59E-05	0.0019479
MGP_FVBNJ_G0024011	Syt4	-2.2326452	0.00269021	0.04130394
MGP_FVBNJ_G0016267	Tmem163	-2.2312515	0.00018973	0.0048114
MGP_FVBNJ_G0034843	Elovl4	-2.226562	0.00169327	0.02898969
MGP_FVBNJ_G0004090	Hnf1aos2	-2.2255129	0.00071666	0.01465598
MGP_FVBNJ_G0028143	Kif12	-2.2176964	8.49E-06	0.00032989
MGP_FVBNJ_G0016376	Lad1	-2.2151959	3.38E-05	0.00111515
MGP_FVBNJ_G0029303	Prom1	-2.2143343	5.25E-07	2.68E-05
MGP_FVBNJ_G0017196	Gstt1	-2.2096816	2.71E-05	0.00092037
MGP_FVBNJ_G0028181	Frem1	-2.2087231	1.44E-05	0.00052919
MGP_FVBNJ_G0029780	Rasal1	-2.2072898	0.0010025	0.01923128
MGP_FVBNJ_G0031065	Plekhg6	-2.2043613	2.58E-05	0.00088638
MGP_FVBNJ_G0030699	Slc4a5	-2.197357	0.00141707	0.02529049
MGP_FVBNJ_G0019785	Esrrb	-2.1970576	0.000618	0.01298345
MGP_FVBNJ_G0026890	Cp	-2.1959888	2.23E-07	1.22E-05
MGP_FVBNJ_G0023427	Ltb	-2.1914546	5.53E-05	0.00168483
MGP_FVBNJ_G0030630	Cd8a	-2.1905348	0.00150108	0.0264301
MGP_FVBNJ_G0025574	Dpp4	-2.1867622	5.21E-06	0.00021192
MGP_FVBNJ_G0025392	NA	-2.1826772	1.29E-07	7.46E-06
MGP_FVBNJ_G0016948	Rspo3	-2.1815743	5.84E-07	2.93E-05
MGP_FVBNJ_G0032739	Aqp8	-2.1745892	0.00010883	0.00296928
MGP_FVBNJ_G0022503	Cldn1	-2.171524	3.82E-05	0.0012383
MGP_FVBNJ_G0016953	Hey2	-2.1676188	2.17E-06	9.64E-05
MGP_FVBNJ_G0015780	Vwa3b	-2.1622331	0.00197401	0.03259463
MGP_FVBNJ_G0028765	Alpl	-2.1617746	2.73E-08	1.76E-06
MGP_FVBNJ_G0027211	Pklr	-2.1610231	0.00017709	0.00452179
MGP_FVBNJ_G0035131	Dclk3	-2.1604981	0.00029184	0.00683352
MGP_FVBNJ_G0033417	Nat1	-2.1496337	0.00214197	0.03472739

MGP_FVBNJ_G0017667	Gpr182	-2.147746	2.36E-05	0.00081563
MGP_FVBNJ_G0016283	C4bp	-2.1434028	2.40E-05	0.00083111
MGP_FVBNJ_G0016990	Mfsd4b3	-2.1429625	5.10E-05	0.00157506
MGP_FVBNJ_G0029452	Ugt2b35	-2.1391117	2.81E-05	0.00094613
MGP_FVBNJ_G0021955	NA	-2.1358536	0.00014474	0.00377809
MGP_FVBNJ_G0035630	NA	-2.1339257	1.17E-12	1.95E-10
MGP_FVBNJ_G0029759	Nos1	-2.130963	0.00233314	0.03722832
MGP_FVBNJ_G0035795	Nap1l3	-2.1290988	0.00012141	0.00324897
MGP_FVBNJ_G0023368	H2-Eb1	-2.1287768	2.25E-05	0.00078385
MGP_FVBNJ_G0035257	Nudt11	-2.1272693	0.00268544	0.04128349
MGP_FVBNJ_G0027833	Esrp1	-2.1245324	0.00113032	0.02110787
MGP_FVBNJ_G0002077	NA	-2.1233241	3.58E-05	0.00117172
MGP_FVBNJ_G0032889	Lhpp	-2.1218667	3.79E-06	0.00015865
MGP_FVBNJ_G0028244	Cyp2j6	-2.1179155	1.14E-08	7.85E-07
MGP_FVBNJ_G0025690	Frzb	-2.1149322	1.95E-08	1.28E-06
MGP_FVBNJ_G0039657	NA	-2.113589	0.00024265	0.00587914
MGP_FVBNJ_G0028516	Heyl	-2.111149	0.00093703	0.01817868
MGP_FVBNJ_G0019862	Asb2	-2.1103559	0.00079166	0.01586958
MGP_FVBNJ_G0015136	Gm15941	-2.108986	0.000199	0.0050123
MGP_FVBNJ_G0025595	B3galt1	-2.102719	0.00019287	0.00488542
MGP_FVBNJ_G0017332	Reep6	-2.1007919	4.61E-05	0.00143903
MGP_FVBNJ_G0029529	Sowahb	-2.1006145	1.63E-09	1.35E-07
MGP_FVBNJ_G0021645	Sybu	-2.0981771	6.00E-06	0.00024028
MGP_FVBNJ_G0027587	Olfm3	-2.0961273	0.00214203	0.03472739
MGP_FVBNJ_G0015513	NA	-2.0957353	0.00017851	0.00454768
MGP_FVBNJ_G0025942	Nr1h3	-2.0946539	5.36E-06	0.00021734
MGP_FVBNJ_G0027391	Car14	-2.0928507	0.00261632	0.04052846
MGP_FVBNJ_G0032848	Itgax	-2.0921709	0.00121661	0.02243138
MGP_FVBNJ_G0028674	Kdf1	-2.0920974	3.92E-05	0.00126163
MGP_FVBNJ_G0033478	Gdf15	-2.0920512	9.10E-05	0.00255689
MGP_FVBNJ_G0021667	Col14a1	-2.0889963	1.84E-06	8.37E-05
MGP_FVBNJ_G0017603	Grip1	-2.0834799	0.00020551	0.00513558
MGP_FVBNJ_G0019248	Foxj1	-2.0817362	0.00029121	0.00682594
MGP_FVBNJ_G0022039	Odf3b	-2.0760537	0.0014252	0.02537466
MGP_FVBNJ_G0015891	Aox3	-2.0712958	0.00017804	0.00454093
MGP_FVBNJ_G0007308	Irx3os	-2.0703789	0.00170798	0.02912968
MGP_FVBNJ_G0026838	Hnf4g	-2.0686138	0.00022936	0.00561219
MGP_FVBNJ_G0025679	Zfp385b	-2.0666523	0.00013924	0.00365572
MGP_FVBNJ_G0033641	Klf1	-2.0656705	0.00269047	0.04130394
MGP_FVBNJ_G0039901	NA	-2.0651041	5.64E-05	0.00171391
MGP_FVBNJ_G0034531	Plet1	-2.0627302	0.00132099	0.02390118

MGP_FVBNJ_G0033339	Pdlim3	-2.0585839	5.54E-06	0.00022423
MGP_FVBNJ_G0015871	Pgap1	-2.0574839	2.63E-13	4.93E-11
MGP_FVBNJ_G0027565	5330417C22Rik	-2.0574674	1.12E-05	0.00042103
MGP_FVBNJ_G0024236	Prdm6	-2.055472	0.00026929	0.00637242
MGP_FVBNJ_G0018459	Slc13a5	-2.054153	0.00030565	0.00711929
MGP_FVBNJ_G0024305	Atp8b1	-2.0536772	3.09E-08	1.96E-06
MGP_FVBNJ_G0017690	Il23a	-2.0529359	0.00295173	0.04436747
MGP_FVBNJ_G0016043	Sgpp2	-2.0517214	0.00057961	0.01223875
MGP_FVBNJ_G0018623	Rab11fip4	-2.0515961	7.88E-07	3.83E-05
MGP_FVBNJ_G0026000	Ehf	-2.0511566	4.25E-05	0.00134263
MGP_FVBNJ_G0035366	Klhl13	-2.0497264	3.25E-13	5.94E-11
MGP_FVBNJ_G0023814	Ehd3	-2.0478822	3.73E-06	0.00015678
MGP_FVBNJ_G0033305	Cldn23	-2.0456246	0.00086227	0.0170247
MGP_FVBNJ_G0018095	Col23a1	-2.0439988	0.001053	0.01992575
MGP_FVBNJ_G0032731	Chp2	-2.0436551	2.29E-05	0.00079514
MGP_FVBNJ_G0023878	Slc3a1	-2.0420112	0.00027544	0.00649726
MGP_FVBNJ_G0020767	Itga1	-2.0368489	4.37E-10	4.09E-08
MGP_FVBNJ_G0016586	Pcp4l1	-2.0342822	0.00178309	0.03022571
MGP_FVBNJ_G0018295	Hs3st3b1	-2.0332242	3.62E-07	1.89E-05
MGP_FVBNJ_G0033206	Atp7b	-2.0329668	5.23E-05	0.00161017
MGP_FVBNJ_G0027641	Ugt8a	-2.0300995	0.00010389	0.00285206
MGP_FVBNJ_G0019779	Mfsd7c	-2.0294421	2.31E-05	0.00080326
MGP_FVBNJ_G0018999	Hsd17b1	-2.0279983	0.00108136	0.0203415
MGP_FVBNJ_G0028135	Bspry	-2.0277223	5.48E-05	0.00167301
MGP_FVBNJ_G0032662	Insc	-2.0248478	2.06E-05	0.00073143
MGP_FVBNJ_G0027158	Fcrl1	-2.0169309	0.00011773	0.00317339
MGP_FVBNJ_G0036226	NA	-2.0160306	0.00252436	0.03946036
MGP_FVBNJ_G0023567	Adgrf1	-2.0156156	0.00025203	0.00605205
MGP_FVBNJ_G0023370	Btnl2	-2.0142021	0.00069552	0.01428916
MGP_FVBNJ_G0022661	NA	-2.0123362	7.46E-05	0.00216332
MGP_FVBNJ_G0025319	Sardh	-2.0119781	4.73E-05	0.00147224
MGP_FVBNJ_G0004961	Gm16548	-2.010176	0.00073601	0.01499679
MGP_FVBNJ_G0032978	Ifitm1	-2.0097362	3.38E-06	0.00014312
MGP_FVBNJ_G0019179	Sdk2	-2.0054653	5.47E-07	2.77E-05
MGP_FVBNJ_G0031835	Fxyd1	-2.0048644	5.27E-05	0.00161661
MGP_FVBNJ_G0030040	Stag3	-2.0034794	0.0021365	0.03472739
MGP_FVBNJ_G0023393	Cfb	-1.9996095	4.15E-05	0.0013209
MGP_FVBNJ_G0020283	Bmp6	-1.9992909	5.20E-05	0.00160205
MGP_FVBNJ_G0025127	Ccdc3	-1.9987584	5.77E-05	0.00174436
MGP_FVBNJ_G0029589	Abcg3	-1.9968378	0.00012751	0.00339326
MGP_FVBNJ_G0023516	H2-M2	-1.9957678	0.00188009	0.03146325

MGP_FVBNJ_G0025451	Cutal	-1.9934736	2.96E-05	0.00098908
MGP_FVBNJ_G0023366	H2-Ab1	-1.993446	8.42E-05	0.00239684
MGP_FVBNJ_G0021683	Klhl38	-1.9927233	0.00281159	0.04269292
MGP_FVBNJ_G0031206	Rerg	-1.9909375	7.31E-09	5.19E-07
MGP_FVBNJ_G0027546	Gstm6	-1.9880529	0.00015401	0.00401527
MGP_FVBNJ_G0018728	Sep-04	-1.9866352	1.15E-08	7.86E-07
MGP_FVBNJ_G0020568	Tppp	-1.9847336	0.00041815	0.0093396
MGP_FVBNJ_G0031123	Klrk1	-1.9817753	0.00038743	0.00874988
MGP_FVBNJ_G0035625	Prrg1	-1.9797468	2.15E-06	9.62E-05
MGP_FVBNJ_G0016623	Casq1	-1.9781275	3.22E-05	0.00106714
MGP_FVBNJ_G0034729	Dapk2	-1.9767076	0.00019416	0.00491253
MGP_FVBNJ_G0035124	NA	-1.9763203	0.00034529	0.00790157
MGP_FVBNJ_G0018325	Ntn1	-1.9746079	1.77E-06	8.07E-05
MGP_FVBNJ_G0021546	Npr3	-1.9741687	7.12E-06	0.0002814
MGP_FVBNJ_G0027887	Epha7	-1.9713325	9.36E-10	8.12E-08
MGP_FVBNJ_G0034064	Gucy1a2	-1.9709342	1.90E-06	8.59E-05
MGP_FVBNJ_G0010634	NA	-1.9665926	0.00067881	0.01398462
MGP_FVBNJ_G0035169	Acaa1b	-1.9660484	0.00040995	0.00918406
MGP_FVBNJ_G0034048	Kcnk1	-1.9630065	0.00043883	0.0097333
MGP_FVBNJ_G0016800	G0s2	-1.9627022	0.00032995	0.00759017
MGP_FVBNJ_G0024417	Cd226	-1.9606302	0.00241014	0.03829258
MGP_FVBNJ_G0023944	Abhd3	-1.9568736	7.72E-05	0.00222665
MGP_FVBNJ_G0031286	Nlrp12	-1.9557751	4.47E-06	0.00018384
MGP_FVBNJ_G0020432	Lect2	-1.9553488	0.000262	0.00622634
MGP_FVBNJ_G0035176	Xylb	-1.9532274	3.64E-05	0.0011858
MGP_FVBNJ_G0019884	NA	-1.9515365	0.00053606	0.01153417
MGP_FVBNJ_G0033058	Shank2	-1.9499807	6.11E-05	0.00182717
MGP_FVBNJ_G0016735	Susd4	-1.9462219	0.0021502	0.03480768
MGP_FVBNJ_G0021594	Rida	-1.9437862	6.32E-05	0.0018759
MGP_FVBNJ_G0023126	Igfals	-1.941285	0.00013444	0.00354662
MGP_FVBNJ_G0018298	Hs3st3a1	-1.9401041	2.99E-06	0.0001278
MGP_FVBNJ_G0001030	NA	-1.9388581	2.64E-05	0.00090142
MGP_FVBNJ_G0016369	Lmod1	-1.9346609	8.08E-06	0.0003148
MGP_FVBNJ_G0035242	Cxcr6	-1.9343528	0.00147028	0.02603185
MGP_FVBNJ_G0029258	Ablim2	-1.932653	0.00017123	0.00439216
MGP_FVBNJ_G0025617	Sp5	-1.9320072	0.00065091	0.0135349
MGP_FVBNJ_G0027799	Lrrc7	-1.9311831	0.00325268	0.04776389
MGP_FVBNJ_G0027035	Med12l	-1.9297526	8.57E-07	4.13E-05
MGP_FVBNJ_G0031116	Clec9a	-1.9282758	2.80E-05	0.00094375
MGP_FVBNJ_G0034762	Aqp9	-1.9268172	5.36E-05	0.00164128
MGP_FVBNJ_G0018441	Scimp	-1.9265145	0.00145951	0.02586161

MGP_FVBNJ_G0039519	NA	-1.9259577	0.00047783	0.01047326
MGP_FVBNJ_G0026908	Tnik	-1.924309	5.66E-08	3.47E-06
MGP_FVBNJ_G0019432	Lpin1	-1.9230996	0.00274273	0.04181826
MGP_FVBNJ_G0010316	NA	-1.9120494	0.00112292	0.02100861
MGP_FVBNJ_G0023367	H2-Aa	-1.9100757	0.00013375	0.0035324
MGP_FVBNJ_G0019946	Gm266	-1.9048321	0.00056134	0.01190557
MGP_FVBNJ_G0033367	Tenm3	-1.9040438	0.00077026	0.01555244
MGP_FVBNJ_G0032051	Saa1	-1.9010143	1.25E-06	5.87E-05
MGP_FVBNJ_G0018681	Wfdc18	-1.8951181	0.00194319	0.03225287
MGP_FVBNJ_G0016704	Coq8a	-1.8923196	0.0002248	0.00552489
MGP_FVBNJ_G0019164	Abca8a	-1.8922948	0.0002104	0.00522848
MGP_FVBNJ_G0028841	Fhad1	-1.8922822	0.00031564	0.00731378
MGP_FVBNJ_G0016776	Batf3	-1.8853198	3.86E-05	0.00125118
MGP_FVBNJ_G0023687	Sult1c2	-1.8853063	0.00166093	0.02853563
MGP_FVBNJ_G0015887	Spats2l	-1.8841294	2.71E-07	1.44E-05
MGP_FVBNJ_G0030323	Plxna4	-1.8830905	1.76E-08	1.17E-06
MGP_FVBNJ_G0015695	Msc	-1.8820653	8.18E-05	0.00234263
MGP_FVBNJ_G0018781	Epn3	-1.8768068	0.0019509	0.03233267
MGP_FVBNJ_G0019882	Serpina3c	-1.8760459	0.00019696	0.00497635
MGP_FVBNJ_G0040332	NA	-1.8757365	0.00193948	0.03223938
MGP_FVBNJ_G0038056	Gm25047	-1.8745484	0.00310024	0.04604089
MGP_FVBNJ_G0028830	Tmem82	-1.8729031	5.72E-05	0.00173463
MGP_FVBNJ_G0015935	Adam23	-1.8623444	0.00078283	0.01574906
MGP_FVBNJ_G0035465	Gpc4	-1.8615487	2.44E-07	1.32E-05
MGP_FVBNJ_G0018454	Pitpnm3	-1.8609846	0.00211075	0.0343957
MGP_FVBNJ_G0022202	Krt79	-1.8590362	0.00253474	0.03959483
MGP_FVBNJ_G0026588	Tgm2	-1.8577002	2.71E-06	0.00011604
MGP_FVBNJ_G0033449	Tm6sf2	-1.8534067	9.53E-05	0.00265153
MGP_FVBNJ_G0019204	Ush1g	-1.8516976	0.0003913	0.00881937
MGP_FVBNJ_G0032285	Dlg2	-1.8470091	9.81E-08	5.78E-06
MGP_FVBNJ_G0028273	Dnajc6	-1.8427885	0.00035789	0.00815717
MGP_FVBNJ_G0023989	Asxl3	-1.8421606	9.32E-05	0.00260433
MGP_FVBNJ_G0013976	NA	-1.8405395	0.00135328	0.02430771
MGP_FVBNJ_G0035214	Fam198a	-1.8389017	0.00112343	0.02100861
MGP_FVBNJ_G0018114	Fstl4	-1.836129	0.00248449	0.03910878
MGP_FVBNJ_G0022670	Cd96	-1.8357375	0.00322317	0.04742414
MGP_FVBNJ_G0021270	Gata4	-1.8335093	4.27E-07	2.20E-05
MGP_FVBNJ_G0023389	NA	-1.8313825	0.0002375	0.00577317
MGP_FVBNJ_G0023928	Mkx	-1.8306479	0.00038215	0.00865693
MGP_FVBNJ_G0029503	Cxcl1	-1.8297444	2.68E-06	0.00011519
MGP_FVBNJ_G0021851	Tst	-1.8294315	0.00010148	0.00280313

MGP_FVBNJ_G0017622	Avpr1a	-1.8238578	0.00111014	0.02083017
MGP_FVBNJ_G0005964	A730036I17Rik	-1.8237924	0.00344312	0.04990287
MGP_FVBNJ_G0028029	Aldh1b1	-1.8236685	4.07E-05	0.00130435
MGP_FVBNJ_G0033695	Irx3	-1.8226376	0.00022852	0.0055978
MGP_FVBNJ_G0034266	Ntm	-1.8222603	2.59E-05	0.00088741
MGP_FVBNJ_G0018424	Vmo1	-1.8191758	0.00324713	0.04771371
MGP_FVBNJ_G0015512	NA	-1.818257	0.00091509	0.01781504
MGP_FVBNJ_G0028497	Col9a2	-1.8168606	0.00205213	0.0336353
MGP_FVBNJ_G0029071	Sema3d	-1.8142751	2.15E-05	0.00075658
MGP_FVBNJ_G0021494	NA	-1.8110008	0.00015858	0.00411996
MGP_FVBNJ_G0025530	Lypd6	-1.8104122	0.00081935	0.01632184
MGP_FVBNJ_G0030415	Gstk1	-1.8095328	0.00025301	0.00605205
MGP_FVBNJ_G0035431	Dcaf12l1	-1.8081886	0.00130856	0.02375338
MGP_FVBNJ_G0019869	Ppp4r4	-1.807914	0.00091376	0.01780467
MGP_FVBNJ_G0027706	Mttp	-1.8033513	0.00040557	0.00911344
MGP_FVBNJ_G0028393	Faah	-1.8028816	0.0001278	0.00339554
MGP_FVBNJ_G0016497	Tnfsf4	-1.7975262	0.00239076	0.03803878
MGP_FVBNJ_G0026873	Car2	-1.7970855	4.44E-06	0.00018315
MGP_FVBNJ_G0027049	P2ry1	-1.7960484	2.45E-08	1.59E-06
MGP_FVBNJ_G0016211	Rnf152	-1.7959153	3.09E-05	0.00102868
MGP_FVBNJ_G0039651	NA	-1.7953819	0.00051577	0.01119467
MGP_FVBNJ_G0035669	Vsig4	-1.7945894	0.00100494	0.01926157
MGP_FVBNJ_G0017863	Adcy1	-1.7939755	0.00111532	0.02090975
MGP_FVBNJ_G0017468	Nr1h4	-1.7921957	0.00020803	0.00518692
MGP_FVBNJ_G0017089	Oit3	-1.786522	0.00192933	0.03211854
MGP_FVBNJ_G0020770	Parp8	-1.7857108	0.00034531	0.00790157
MGP_FVBNJ_G0015402	Gm830	-1.7852085	0.0011558	0.0214701
MGP_FVBNJ_G0025634	Rapgef4	-1.7843176	0.00025604	0.0061107
MGP_FVBNJ_G0016673	Rgs7	-1.7843065	0.00028104	0.00661525
MGP_FVBNJ_G0020314	Cd83	-1.7758135	0.00030723	0.00714868
MGP_FVBNJ_G0027986	Cd72	-1.7754273	0.00158491	0.02757949
MGP_FVBNJ_G0036783	Gm13053	-1.7741179	8.38E-05	0.00239088
MGP_FVBNJ_G0033938	Osgin1	-1.7733614	0.00042293	0.00942757
MGP_FVBNJ_G0020797	Acox2	-1.7714974	0.00256845	0.03995339
MGP_FVBNJ_G0026591	Lbp	-1.7693497	6.23E-07	3.09E-05
MGP_FVBNJ_G0020393	Unc5a	-1.7687893	0.00134249	0.02416136
MGP_FVBNJ_G0039551	NA	-1.7647007	0.00013889	0.00365105
MGP_FVBNJ_G0003291	Gm15638	-1.7644315	0.00011168	0.00303588
MGP_FVBNJ_G0024653	Tkfc	-1.7640223	0.00249479	0.03916285
MGP_FVBNJ_G0023671	Unc5cl	-1.7622434	0.00023665	0.00575896
MGP_FVBNJ_G0035980	Reps2	-1.762126	0.00018361	0.0046721

MGP_FVBNJ_G0016628	Kcnj10	-1.7613341	0.00049858	0.010864
MGP_FVBNJ_G0009866	Gbp11	-1.7600106	0.00133305	0.02406074
MGP_FVBNJ_G0032570	Olfml1	-1.7583718	0.00030514	0.00711484
MGP_FVBNJ_G0022211	Csad	-1.7559476	7.54E-05	0.002184
MGP_FVBNJ_G0033233	1810011O10Rik	-1.7550194	1.14E-08	7.85E-07
MGP_FVBNJ_G0030589	Nap1l5	-1.7533193	0.00062746	0.0131207
MGP_FVBNJ_G0023693	Adgre4	-1.753299	0.00076834	0.01552762
MGP_FVBNJ_G0016564	Rgs5	-1.7531678	7.29E-05	0.00212981
MGP_FVBNJ_G0009946	NA	-1.7525208	0.00117282	0.02173179
MGP_FVBNJ_G0018416	Alox12e	-1.7472026	0.00102687	0.01954729
MGP_FVBNJ_G0021596	Nipal2	-1.7459206	0.00012537	0.00334683
MGP_FVBNJ_G0024839	Prkg1	-1.7440115	7.67E-11	8.72E-09
MGP_FVBNJ_G0035152	Cmtm8	-1.7438033	0.00026105	0.00621036
MGP_FVBNJ_G0018422	Zmynd15	-1.7426604	0.00016235	0.00420116
MGP_FVBNJ_G0029761	Tesc	-1.741811	3.53E-05	0.00115833
MGP_FVBNJ_G0025683	Cerkl	-1.7373715	0.00021531	0.00532086
MGP_FVBNJ_G0017162	Tmem26	-1.736981	8.66E-06	0.00033384
MGP_FVBNJ_G0016281	Cd55	-1.7348274	0.00011657	0.00314571
MGP_FVBNJ_G0023281	Tff2	-1.730647	0.00107441	0.0202326
MGP_FVBNJ_G0022599	Ildr1	-1.7295061	0.00078658	0.01579737
MGP_FVBNJ_G0002946	Dnah2os	-1.7287436	4.34E-05	0.00136757
MGP_FVBNJ_G0019883	Serpina3i	-1.7250691	2.18E-05	0.00076557
MGP_FVBNJ_G0031025	NA	-1.7250048	0.00041099	0.00919823
MGP_FVBNJ_G0028625	Serinc2	-1.724446	1.02E-05	0.00038564
MGP_FVBNJ_G0007130	1810008I18Rik	-1.7229692	0.00125411	0.0229518
MGP_FVBNJ_G0035246	Ccr3	-1.721508	0.00251073	0.03932078
MGP_FVBNJ_G0003718	Gm15998	-1.7192833	0.00075509	0.01530146
MGP_FVBNJ_G0031024	NA	-1.7145455	1.67E-05	0.0006071
MGP_FVBNJ_G0002064	NA	-1.7125782	1.23E-07	7.15E-06
MGP_FVBNJ_G0032865	Fgfr2	-1.7110254	1.99E-07	1.10E-05
MGP_FVBNJ_G0029782	Oas2	-1.7087209	0.0030403	0.04530168
MGP_FVBNJ_G0020695	Marveld2	-1.7069621	0.00012757	0.00339326
MGP_FVBNJ_G0016342	Prelp	-1.704888	0.00136218	0.02440869
MGP_FVBNJ_G0017186	Upb1	-1.7029996	0.00055956	0.01188274
MGP_FVBNJ_G0027723	Gbp5	-1.7024355	0.00244231	0.03858639
MGP_FVBNJ_G0002586	Gm9947	-1.7014726	0.00088017	0.01729083
MGP_FVBNJ_G0032213	Homer2	-1.7006077	0.00192348	0.03206904
MGP_FVBNJ_G0028238	Hook1	-1.6993644	0.00086795	0.01712158
MGP_FVBNJ_G0022571	NA	-1.6967666	1.27E-06	5.94E-05
MGP_FVBNJ_G0027725	NA	-1.6966183	6.78E-05	0.00199364
MGP_FVBNJ_G0033450	Hapln4	-1.6942793	1.51E-05	0.00055348

MGP_FVBNJ_G0019258	Cygb	-1.6906266	8.05E-05	0.00231267
MGP_FVBNJ_G0023886	Epas1	-1.6892215	1.84E-07	1.02E-05
MGP_FVBNJ_G0026719	Fam65c	-1.6883548	3.99E-05	0.00128046
MGP_FVBNJ_G0030880	Cav3	-1.6877079	0.00169743	0.02903871
MGP_FVBNJ_G0010771	NA	-1.6870941	0.00117776	0.02180517
MGP_FVBNJ_G0026845	Il7	-1.6848082	0.0022861	0.03658251
MGP_FVBNJ_G0032797	Qprt	-1.6839756	0.00078735	0.01579737
MGP_FVBNJ_G0019906	Ak7	-1.6834564	0.00052354	0.01133409
MGP_FVBNJ_G0023438	H2-Q10	-1.6832049	0.0002493	0.00602065
MGP_FVBNJ_G0031801	Prodh2	-1.6829825	0.00024108	0.00584748
MGP_FVBNJ_G0022574	Mylk	-1.6814406	1.31E-07	7.56E-06
MGP_FVBNJ_G0023136	Tmem204	-1.677786	1.18E-06	5.56E-05
MGP_FVBNJ_G0032150	Fam174b	-1.6770027	0.00102286	0.01950429
MGP_FVBNJ_G0028973	NA	-1.6741936	0.00191643	0.0319754
MGP_FVBNJ_G0030331	Akr1b7	-1.6733202	0.00195531	0.03235768
MGP_FVBNJ_G0025414	Slc25a25	-1.6674327	0.00052525	0.01134998
MGP_FVBNJ_G0030203	Ppp1r9a	-1.6639696	3.24E-05	0.00107108
MGP_FVBNJ_G0018674	Ccl5	-1.6596691	0.00199092	0.03282517
MGP_FVBNJ_G0020038	Chrm3	-1.6590519	0.00211395	0.0343957
MGP_FVBNJ_G0022257	Mefv	-1.6573547	0.00085855	0.01696633
MGP_FVBNJ_G0022352	Myh11	-1.6556817	0.00123204	0.0226409
MGP_FVBNJ_G0022099	Zfp641	-1.6546948	0.00015745	0.00410013
MGP_FVBNJ_G0029509	Btc	-1.6529537	0.00128422	0.02336865
MGP_FVBNJ_G0016280	Cd55b	-1.6512026	0.00025066	0.00602817
MGP_FVBNJ_G0016006	Cyp27a1	-1.6506312	0.00047054	0.01034407
MGP_FVBNJ_G0024869	Pank1	-1.6494853	0.00143943	0.02556689
MGP_FVBNJ_G0025985	Lrrc4c	-1.6491035	0.00024572	0.00594715
MGP_FVBNJ_G0028100	Palm2	-1.6490202	0.00082271	0.01635956
MGP_FVBNJ_G0033566	Ednra	-1.6487814	1.03E-06	4.90E-05
MGP_FVBNJ_G0026219	Slc24a5	-1.6482396	0.0015422	0.02702601
MGP_FVBNJ_G0019888	Serpina3n	-1.6480526	9.67E-05	0.00268644
MGP_FVBNJ_G0029025	AW011738	-1.6440843	2.28E-06	0.00010041
MGP_FVBNJ_G0028842	Tmem51	-1.6429149	1.54E-07	8.72E-06
MGP_FVBNJ_G0035274	Magix	-1.638915	0.00279631	0.04255507
MGP_FVBNJ_G0032807	Itgal	-1.6386667	0.00025873	0.00616167
MGP_FVBNJ_G0031564	Apoe	-1.6374853	0.00034603	0.00790297
MGP_FVBNJ_G0018705	Tbx2	-1.6360162	2.68E-07	1.43E-05
MGP_FVBNJ_G0010085	NA	-1.6350113	0.00193418	0.03217515
MGP_FVBNJ_G0033928	Bco1	-1.6336783	0.00019786	0.00498915
MGP_FVBNJ_G0023425	Aif1	-1.6305825	0.0022823	0.03654793
MGP_FVBNJ_G0031026	NA	-1.6292302	0.00028204	0.00663178

MGP_FVBNJ_G0022191	Krt75	-1.6275139	0.00247724	0.0390248
MGP_FVBNJ_G0026258	Kcnip3	-1.6269443	6.13E-05	0.00183064
MGP_FVBNJ_G0029522	Cxcl10	-1.6249993	0.0025119	0.03932078
MGP_FVBNJ_G0025979	NA	-1.6245824	8.65E-06	0.00033384
MGP_FVBNJ_G0026861	Slc10a5	-1.6233125	0.00304684	0.04536887
MGP_FVBNJ_G0020993	Lrit1	-1.6231615	0.00341046	0.0494616
MGP_FVBNJ_G0016604	Tstd1	-1.6219579	0.00048001	0.01051074
MGP_FVBNJ_G0021614	Ncald	-1.621186	2.51E-06	0.00010828
MGP_FVBNJ_G0022495	Rtp4	-1.6202438	0.00054568	0.01167351
MGP_FVBNJ_G0024257	Chsy3	-1.6178853	7.18E-05	0.00210699
MGP_FVBNJ_G0017878	Ddc	-1.6161244	0.00156559	0.0273286
MGP_FVBNJ_G0018476	Itgae	-1.6120126	0.00333555	0.04859652
MGP_FVBNJ_G0015559	AV026068	-1.6099704	1.68E-08	1.12E-06
MGP_FVBNJ_G0018972	Hap1	-1.6059518	2.94E-05	0.0009826
MGP_FVBNJ_G0027653	Enpep	-1.6030001	0.00082267	0.01635956
MGP_FVBNJ_G0031275	Tmtc1	-1.6017668	0.0003905	0.00881026
MGP_FVBNJ_G0030063	Tmem184a	-1.6013313	0.00121059	0.02237578
MGP_FVBNJ_G0024063	Slc23a1	-1.6004654	0.0007072	0.01448916
MGP_FVBNJ_G0019202	Fads6	-1.5955495	0.00040795	0.0091578
MGP_FVBNJ_G0021911	Grap2	-1.5954952	0.00083205	0.01650115
MGP_FVBNJ_G0031223	Pde3a	-1.5924138	6.05E-07	3.01E-05
MGP_FVBNJ_G0015607	Gm17396	-1.591883	8.75E-05	0.00247847
MGP_FVBNJ_G0017498	Plxnc1	-1.5880054	7.40E-05	0.00214957
MGP_FVBNJ_G0021953	Cyp2d22	-1.5861422	0.00050364	0.01095282
MGP_FVBNJ_G0019261	St6galnac2	-1.584084	6.90E-06	0.00027382
MGP_FVBNJ_G0030467	Rarres2	-1.5836175	0.00044637	0.00989059
MGP_FVBNJ_G0026107	Rasgrp1	-1.5816469	0.0008804	0.01729083
MGP_FVBNJ_G0024876	Ppp1r3c	-1.5786324	0.00291802	0.04403891
MGP_FVBNJ_G0022897	Bace2	-1.5776149	3.32E-10	3.22E-08
MGP_FVBNJ_G0027764	Adgrl2	-1.5746649	2.79E-09	2.19E-07
MGP_FVBNJ_G0025123	Phyh	-1.5734862	0.00214702	0.03478149
MGP_FVBNJ_G0033626	Cacna1a	-1.5728158	0.00025317	0.00605205
MGP_FVBNJ_G0024268	Cd74	-1.5726831	0.00107467	0.0202326
MGP_FVBNJ_G0025581	Grb14	-1.5721366	0.00021472	0.0053123
MGP_FVBNJ_G0019727	Dpf3	-1.5710164	0.00309986	0.04604089
MGP_FVBNJ_G0004681	Tsix	-1.5693809	0.00055819	0.01187267
MGP_FVBNJ_G0020447	NA	-1.5689773	0.00087932	0.01729083
MGP_FVBNJ_G0024770	Pcsk5	-1.5668194	3.71E-07	1.93E-05
MGP_FVBNJ_G0027837	Gem	-1.5655365	1.91E-05	0.00068251
MGP_FVBNJ_G0017223	Col18a1	-1.5654378	0.00026618	0.00630549
MGP_FVBNJ_G0034446	C1qtnf5	-1.5628168	0.00075585	0.01530302

MGP_FVBNJ_G0027825	Gdf6	-1.5579303	0.0025845	0.04017515
MGP_FVBNJ_G0017883	Egfr	-1.5531176	7.20E-05	0.00210763
MGP_FVBNJ_G0017148	Dnajc12	-1.5515673	0.00048986	0.01070536
MGP_FVBNJ_G0032571	Ppfibp2	-1.5507085	1.51E-05	0.00055348
MGP_FVBNJ_G0002078	NA	-1.5506593	0.00105737	0.01997447
MGP_FVBNJ_G0028329	Scp2	-1.5495192	0.00052036	0.01128341
MGP_FVBNJ_G0030080	Card11	-1.5474307	0.00280385	0.04260445
MGP_FVBNJ_G0021515	Lifr	-1.5437552	0.00011914	0.00319593
MGP_FVBNJ_G0031240	Abcc9	-1.5408368	5.02E-06	0.00020513
MGP_FVBNJ_G0035979	Nhs	-1.539863	0.00037019	0.00841176
MGP_FVBNJ_G0015588	Gm16551	-1.5358482	0.00069845	0.01433625
MGP_FVBNJ_G0022968	Smoc2	-1.5354479	0.00303126	0.04519719
MGP_FVBNJ_G0029366	Limch1	-1.5343896	2.68E-05	0.00090976
MGP_FVBNJ_G0028802	Klhdc7a	-1.5338591	0.00279719	0.04255507
MGP_FVBNJ_G0028676	Nr0b2	-1.533525	0.00255522	0.03985893
MGP_FVBNJ_G0024116	Pcdhb8	-1.5335079	2.19E-05	0.00076557
MGP_FVBNJ_G0019901	D430019H16Rik	-1.5312349	0.00311452	0.04619149
MGP_FVBNJ_G0027448	Hao2	-1.5296925	0.00016015	0.00415584
MGP_FVBNJ_G0032942	Sprn	-1.5246866	0.00124864	0.02288932
MGP_FVBNJ_G0022664	BC016579	-1.5232337	0.00054135	0.01160307
MGP_FVBNJ_G0021973	Bik	-1.5205153	0.00080321	0.01604331
MGP_FVBNJ_G0034745	Rora	-1.51997	3.63E-05	0.00118317
MGP_FVBNJ_G0031561	Gm44805	-1.5193038	0.00219833	0.03535565
MGP_FVBNJ_G0031644	NA	-1.5170097	0.00089104	0.01740761
MGP_FVBNJ_G0020458	Dapk1	-1.5144089	1.71E-06	7.86E-05
MGP_FVBNJ_G0029457	Ugt2a1	-1.5115922	0.00329322	0.0481056
MGP_FVBNJ_G0032371	Il18bp	-1.5114884	0.00011611	0.00313723
MGP_FVBNJ_G0021725	Khdrbs3	-1.5100444	0.00054897	0.01173268
MGP_FVBNJ_G0019360	Sectm1a	-1.5077392	0.00012167	0.00325187
MGP_FVBNJ_G0027219	Muc1	-1.5044181	0.00046487	0.01023961
MGP_FVBNJ_G0027224	Efna1	-1.5015405	0.00158954	0.02761679
MGP_FVBNJ_G0024645	Syt7	-1.4996031	0.00286311	0.04335711
MGP_FVBNJ_G0025060	Gucy2g	-1.494459	0.00170598	0.02911785
MGP_FVBNJ_G0021157	Lrrc16b	-1.4921822	0.00021137	0.00524693
MGP_FVBNJ_G0016256	Insig2	-1.4901132	0.00248783	0.03910878
MGP_FVBNJ_G0034966	Acad11	-1.4886144	0.00192917	0.03211854
MGP_FVBNJ_G0015243	Gm13415	-1.4864957	0.00336615	0.04894638
MGP_FVBNJ_G0035335	Maob	-1.4844294	0.00287411	0.04346465
MGP_FVBNJ_G0017613	Tbc1d30	-1.4833712	2.08E-05	0.00073461
MGP_FVBNJ_G0022168	Smagp	-1.4830961	0.00031299	0.0072601
MGP_FVBNJ_G0032178	Anpep	-1.481592	4.94E-05	0.0015299

MGP_FVBNJ_G0036772	1700012D16Rik	-1.4815819	0.00166113	0.02853563
MGP_FVBNJ_G0022131	Dnajc22	-1.4794928	0.00160247	0.02773351
MGP_FVBNJ_G0030537	Ggct	-1.4788705	0.00105454	0.01993793
MGP_FVBNJ_G0032750	Sbk1	-1.4733015	0.00175301	0.02982935
MGP_FVBNJ_G0024551	Batf2	-1.472789	0.00331307	0.04830058
MGP_FVBNJ_G0024360	Lipg	-1.4666614	0.00155685	0.02723999
MGP_FVBNJ_G0021774	Naprt	-1.4659563	0.00165396	0.02849178
MGP_FVBNJ_G0016522	Sele	-1.4623807	0.00324564	0.04771371
MGP_FVBNJ_G0033873	Marveld3	-1.4593948	0.00016476	0.00424577
MGP_FVBNJ_G0031760	Ppp1r14a	-1.4482284	0.00293961	0.04426075
MGP_FVBNJ_G0019640	Rtn1	-1.4468227	0.00052547	0.01134998
MGP_FVBNJ_G0030298	Kcp	-1.4445265	0.00149236	0.02631819
MGP_FVBNJ_G0015763	Hs6st1	-1.4412305	8.59E-05	0.00244061
MGP_FVBNJ_G0002065	NA	-1.4385185	6.31E-05	0.001874
MGP_FVBNJ_G0026791	Slco4a1	-1.4384003	0.00296987	0.04455002
MGP_FVBNJ_G0027846	Runx1t1	-1.4376416	7.34E-05	0.00213816
MGP_FVBNJ_G0017510	Lum	-1.4368986	0.00026257	0.00623326
MGP_FVBNJ_G0019518	Agmo	-1.432307	0.0018695	0.03140398
MGP_FVBNJ_G0025687	Pde1a	-1.4288873	5.11E-10	4.67E-08
MGP_FVBNJ_G0026255	Adra2b	-1.4265656	0.00294264	0.04426075
MGP_FVBNJ_G0032732	Prkcb	-1.4265172	0.00299309	0.04479223
MGP_FVBNJ_G0017168	Slc16a9	-1.4174359	0.00048166	0.01053652
MGP_FVBNJ_G0019419	Kcns3	-1.4174295	0.00022651	0.00556068
MGP_FVBNJ_G0034222	Tmem205	-1.4159597	0.00176274	0.02994906
MGP_FVBNJ_G0030002	Tfr2	-1.4155508	0.00296181	0.04445914
MGP_FVBNJ_G0019949	Ckb	-1.4112572	0.00129039	0.02344252
MGP_FVBNJ_G0021275	Xkr6	-1.4102227	0.00025225	0.00605205
MGP_FVBNJ_G0028493	Rims3	-1.4099487	0.00182456	0.03083493
MGP_FVBNJ_G0031857	Kctd15	-1.4075077	0.00011833	0.00318183
MGP_FVBNJ_G0040204	NA	-1.4074453	0.00064831	0.01349336
MGP_FVBNJ_G0030508	Hoxa1	-1.4066094	0.00016265	0.00420116
MGP_FVBNJ_G0032806	NA	-1.4051202	0.00179665	0.03040944
MGP_FVBNJ_G0018036	NA	-1.4016186	0.00255707	0.03985993
MGP_FVBNJ_G0015801	Npas2	-1.3983215	0.00045271	0.00998162
MGP_FVBNJ_G0034271	St14	-1.3930707	0.00068137	0.01402436
MGP_FVBNJ_G0018061	Flt4	-1.381117	5.10E-05	0.00157506
MGP_FVBNJ_G0003935	Tmem51os1	-1.3710638	0.0001301	0.00345258
MGP_FVBNJ_G0022518	Hes1	-1.3668108	2.65E-07	1.42E-05
MGP_FVBNJ_G0021156	Dhrs4	-1.3604166	0.00074574	0.01513477
MGP_FVBNJ_G0023388	Stk19	-1.3575675	0.00264234	0.04081796
MGP_FVBNJ_G0030197	Gng11	-1.3533608	0.00142114	0.02532268

MGP_FVBNJ_G0015811	Il1r1	-1.3521619	3.62E-08	2.30E-06
MGP_FVBNJ_G0018529	Rtn4rl1	-1.3463263	0.00059878	0.01261578
MGP_FVBNJ_G0021307	Adra1a	-1.3459231	0.00328876	0.04807204
MGP_FVBNJ_G0019425	Fam49a	-1.3432648	7.06E-06	0.00027972
MGP_FVBNJ_G0032299	Thrsp	-1.3393086	0.00103915	0.01971004
MGP_FVBNJ_G0035154	Osbpl10	-1.3385437	5.95E-05	0.00178872
MGP_FVBNJ_G0035607	F8	-1.3366859	0.00103878	0.01971004
MGP_FVBNJ_G0031453	Sult2a6	-1.3363547	0.00044948	0.00993976
MGP_FVBNJ_G0021443	Slain1	-1.3341211	0.0024366	0.03857548
MGP_FVBNJ_G0023841	Qpct	-1.3330622	1.78E-06	8.11E-05
MGP_FVBNJ_G0017443	Syn3	-1.330571	0.00032846	0.00756371
MGP_FVBNJ_G0033323	Mtus1	-1.3246645	0.00153181	0.02688625
MGP_FVBNJ_G0022893	Igsf5	-1.3219349	0.00249188	0.03914485
MGP_FVBNJ_G0016310	Slc45a3	-1.317578	0.00051309	0.01114749
MGP_FVBNJ_G0024864	NA	-1.3158715	0.00205161	0.0336353
MGP_FVBNJ_G0020653	Iqgap2	-1.3153197	0.00262802	0.04065316
MGP_FVBNJ_G0018082	Zfp454	-1.3112292	0.00112227	0.02100861
MGP_FVBNJ_G0033686	Sall1	-1.3109727	0.00029625	0.00692939
MGP_FVBNJ_G0002008	NA	-1.3088242	0.00224208	0.03599129
MGP_FVBNJ_G0031246	Sox5	-1.3087336	0.00036372	0.00827306
MGP_FVBNJ_G0018421	Cxcl16	-1.3085872	0.00259527	0.04028631
MGP_FVBNJ_G0016344	Btg2	-1.2984997	0.00083989	0.01661223
MGP_FVBNJ_G0023575	Rcan2	-1.2856058	0.00020858	0.00519481
MGP_FVBNJ_G0034184	A230050P20Rik	-1.2841451	0.00049547	0.0108107
MGP_FVBNJ_G0022210	NA	-1.2826036	0.00083623	0.01656928
MGP_FVBNJ_G0016573	Nos1ap	-1.2791933	0.00337646	0.04906437
MGP_FVBNJ_G0023915	Bambi	-1.2755166	0.00025795	0.00614975
MGP_FVBNJ_G0022465	Thpo	-1.2748567	0.00299405	0.04479223
MGP_FVBNJ_G0035558	Bgn	-1.2733027	0.00259941	0.04030081
MGP_FVBNJ_G0034786	Wdr72	-1.2723003	0.00168749	0.028913
MGP_FVBNJ_G0024685	Ms4a4d	-1.2716574	0.00022693	0.00556484
MGP_FVBNJ_G0001061	NA	-1.2635702	0.00032371	0.00747749
MGP_FVBNJ_G0024926	Pik3ap1	-1.2610638	9.28E-05	0.00259854
MGP_FVBNJ_G0022174	Fignl2	-1.2584297	3.80E-05	0.00123605
MGP_FVBNJ_G0028330	Echdc2	-1.2578439	0.00094013	0.01820267
MGP_FVBNJ_G0035628	Tmem47	-1.2572104	5.54E-05	0.00168582
MGP_FVBNJ_G0018019	Adam19	-1.2552149	0.00140324	0.0250638
MGP_FVBNJ_G0024386	Setbp1	-1.2453994	2.52E-05	0.0008675
MGP_FVBNJ_G0016993	Slc16a10	-1.2358277	0.00283854	0.04304349
MGP_FVBNJ_G0023986	Garem	-1.2334947	0.00020739	0.00517696
MGP_FVBNJ_G0029758	Ksr2	-1.2330126	0.00158805	0.02761243

MGP_FVBNJ_G0028843	NA	-1.2258015	8.77E-08	5.20E-06
MGP_FVBNJ_G0019453	Grhl1	-1.2219368	0.0001848	0.00469322
MGP_FVBNJ_G0021773	Mroh6	-1.2096837	0.00088247	0.01731622
MGP_FVBNJ_G0031518	Pglyrp1	-1.209556	0.00211259	0.0343957
MGP_FVBNJ_G0030978	Cecr2	-1.1982041	0.00202307	0.03324805
MGP_FVBNJ_G0022640	Zdhhc23	-1.196767	0.00272463	0.04165622
MGP_FVBNJ_G0033309	Dlc1	-1.1896313	9.44E-07	4.50E-05
MGP_FVBNJ_G0027520	Dennd2d	-1.1894526	0.00095648	0.01845971
MGP_FVBNJ_G0029682	Adrbk2	-1.1857119	0.00148929	0.02630563
MGP_FVBNJ_G0029527	Ccdc158	-1.1849699	0.00115481	0.0214701
MGP_FVBNJ_G0017722	NA	-1.1845985	0.00335891	0.04890491
MGP_FVBNJ_G0019509	Tspan13	-1.1843756	0.00131764	0.02385994
MGP_FVBNJ_G0031863	Lrp3	-1.1791323	0.00232418	0.0371385
MGP_FVBNJ_G0023897	Epcam	-1.1785766	0.00210288	0.03434086
MGP_FVBNJ_G0030201	Sgce	-1.1776347	0.00265407	0.04091411
MGP_FVBNJ_G0028417	Zswim5	-1.1732789	0.00032548	0.00751067
MGP_FVBNJ_G0022296	AU021092	-1.1718256	0.0033931	0.04927398
MGP_FVBNJ_G0028280	Pde4b	-1.1678472	0.00023565	0.00574088
MGP_FVBNJ_G0021479	Gpr183	-1.1668118	0.00056746	0.01202403
MGP_FVBNJ_G0024069	Ecscr	-1.1593703	0.0032636	0.04789256
MGP_FVBNJ_G0035158	Rbms3	-1.1588181	0.00038407	0.00868276
MGP_FVBNJ_G0028296	Plpp3	-1.1534302	0.00081643	0.01627821
MGP_FVBNJ_G0001436	NA	-1.1467148	0.00072537	0.01479343
MGP_FVBNJ_G0030923	Mkrn2os	-1.1453764	0.00142957	0.02543216
MGP_FVBNJ_G0025715	Serping1	-1.1429977	0.00121416	0.02240472
MGP_FVBNJ_G0020960	Vstm4	-1.140144	0.00244196	0.03858639
MGP_FVBNJ_G0016637	Dusp23	-1.1389709	0.00157215	0.02741279
MGP_FVBNJ_G0001437	NA	-1.1373362	0.00115847	0.02150156
MGP_FVBNJ_G0028236	Jun	-1.1359503	0.00089034	0.01740761
MGP_FVBNJ_G0025423	Eng	-1.1348209	0.00194305	0.03225287
MGP_FVBNJ_G0017509	Dcn	-1.1307339	0.00011323	0.00306317
MGP_FVBNJ_G0028050	Tgfbr1	-1.1301687	4.82E-07	2.48E-05
MGP_FVBNJ_G0019411	Sdc1	-1.1269809	0.00079565	0.01590648
MGP_FVBNJ_G0035825	Armcx6	-1.1236158	0.00139285	0.02489821
MGP_FVBNJ_G0016626	Atp1a2	-1.1132455	0.00133606	0.0240955
MGP_FVBNJ_G0035164	Itga9	-1.105215	3.77E-06	0.00015831
MGP_FVBNJ_G0022506	Il1rap	-1.104983	0.00273885	0.04179475
MGP_FVBNJ_G0021723	St3gal1	-1.1008851	0.00046578	0.01024957
MGP_FVBNJ_G0030207	Asb4	-1.0990721	0.00226549	0.03633103
MGP_FVBNJ_G0019944	Exoc3l4	-1.0925037	0.00101996	0.0194656
MGP_FVBNJ_G0020328	Nhlrc1	-1.0867089	0.00115526	0.0214701

MGP_FVBNJ_G0004910	F730311O21Rik	-1.0856897	0.00023253	0.00568345
MGP_FVBNJ_G0020720	Rgs7bp	-1.0781281	0.00067711	0.01397552
MGP_FVBNJ_G0015923	Raph1	-1.0760976	2.00E-05	0.00070989
MGP_FVBNJ_G0032310	Omp	-1.0693471	0.00124733	0.02288415
MGP_FVBNJ_G0034043	Sipa1l2	-1.0604728	0.00147261	0.02605235
MGP_FVBNJ_G0030827	Fgd5	-1.0573112	4.41E-05	0.00138604
MGP_FVBNJ_G0031791	Syne4	-1.0435057	6.55E-05	0.00194046
MGP_FVBNJ_G0001574	NA	-1.0434141	0.00133859	0.02412167
MGP_FVBNJ_G0029305	Ldb2	-1.0420908	0.0007586	0.01534483
MGP_FVBNJ_G0030832	Prickle2	-1.0317701	7.90E-05	0.00227343
MGP_FVBNJ_G0024072	Cxxc5	-1.0181664	0.00307855	0.04581036
MGP_FVBNJ_G0002103	NA	-1.0139052	0.00302626	0.04515291
MGP_FVBNJ_G0024132	Pcdhga1	-1.0089462	0.00087219	0.01717484
MGP_FVBNJ_G0015908	Mpp4	-1.0079916	0.00198005	0.03267016
MGP_FVBNJ_G0027852	Tmem64	-1.0055861	9.17E-05	0.00257106
MGP_FVBNJ_G0018829	Hoxb4	-1.0039279	0.00017042	0.00438166
MGP_FVBNJ_G0015658	Sox17	-1.0015202	0.00126408	0.02307753
MGP_FVBNJ_G0011307	NA	-1.0012984	0.00113063	0.02110787
MGP_FVBNJ_G0016786	Slc30a1	-1.0008855	0.00232896	0.03718825
MGP_FVBNJ_G0030540	Inmt	-0.9942106	0.00011243	0.00304881
MGP_FVBNJ_G0016433	Rgl1	-0.9865909	8.61E-05	0.00244306
MGP_FVBNJ_G0019866	Ifi27	-0.9865317	0.00199838	0.03291803
MGP_FVBNJ_G0031220	Plekha5	-0.9789148	0.001035	0.01966847
MGP_FVBNJ_G0032483	NA	-0.9747602	0.00062816	0.0131231
MGP_FVBNJ_G0028483	Hivep3	-0.9715993	0.00024891	0.00601775
MGP_FVBNJ_G0022688	Alcam	-0.9625816	0.00070971	0.01452725
MGP_FVBNJ_G0021346	Slc39a14	-0.9549519	0.00207088	0.03386801
MGP_FVBNJ_G0021682	Fbxo32	-0.9528517	0.00106478	0.02008036
MGP_FVBNJ_G0027136	Trim2	-0.9509717	2.92E-05	0.00097933
MGP_FVBNJ_G0029721	Oasl1	-0.9494705	0.00195408	0.03235768
MGP_FVBNJ_G0024216	Sema6a	-0.9460074	0.00021661	0.00534124
MGP_FVBNJ_G0037293	Gm22107	-0.9452824	0.00155206	0.02717755
MGP_FVBNJ_G0030301	Tspan33	-0.9175242	0.00017067	0.0043829
MGP_FVBNJ_G0002101	NA	-0.9162361	0.00055419	0.01182145
MGP_FVBNJ_G0023783	Ptprm	-0.9158988	0.00267747	0.04124633
MGP_FVBNJ_G0007245	E230013L22Rik	-0.9099908	0.00167673	0.02877281
MGP_FVBNJ_G0016653	NA	-0.9099255	0.00288482	0.04359691
MGP_FVBNJ_G0020174	Hfe	-0.9090905	0.00052689	0.01136967
MGP_FVBNJ_G0019945	Tnfaip2	-0.9068229	0.0021128	0.0343957
MGP_FVBNJ_G0022560	Iqcg	-0.9032907	0.00219063	0.03530821
MGP_FVBNJ_G0028586	Trim62	-0.9027369	9.79E-06	0.00037367

MGP_FVBNJ_G0015927	Pard3b	-0.9010212	0.00026533	0.00629194
MGP_FVBNJ_G0022784	Adamts5	-0.8824613	0.00291038	0.04395345
MGP_FVBNJ_G0031874	Tshz3	-0.8809177	0.00185296	0.03117308
MGP_FVBNJ_G0002102	NA	-0.8782074	0.00049565	0.0108107
MGP_FVBNJ_G0017065	Man1a	-0.8635094	0.00271115	0.04150706
MGP_FVBNJ_G0029393	Tec	-0.8573019	0.00184151	0.03107951
MGP_FVBNJ_G0015909	Als2	-0.853987	0.00186311	0.0313202
MGP_FVBNJ_G0022578	Pdia5	-0.8536589	0.00302222	0.04512275
MGP_FVBNJ_G0019161	Fam20a	-0.8510466	0.00285249	0.04322555
MGP_FVBNJ_G0028162	Megf9	-0.8505881	0.00273546	0.04179304
MGP_FVBNJ_G0030873	Itpr1	-0.8482045	0.0003031	0.00707479
MGP_FVBNJ_G0024123	Pcdhb16	-0.8285987	0.00025049	0.00602817
MGP_FVBNJ_G0027648	Alpk1	-0.8210548	0.00315502	0.04663703
MGP_FVBNJ_G0024125	Pcdhb18	-0.8136162	0.00340951	0.0494616
MGP_FVBNJ_G0022357	Snai2	-0.8102366	0.00177172	0.03005588
MGP_FVBNJ_G0019142	Pitpnc1	-0.8090541	0.00317691	0.0468673
MGP_FVBNJ_G0023936	Colec12	-0.8084028	0.00227286	0.03642288
MGP_FVBNJ_G0021697	Trib1	-0.8078548	7.75E-06	0.0003041
MGP_FVBNJ_G0016424	Ivns1abp	-0.8077673	9.20E-06	0.0003528
MGP_FVBNJ_G0016423	Hmcn1	-0.8052148	0.00298074	0.04468303
MGP_FVBNJ_G0032619	St5	-0.8018243	8.39E-05	0.00239088
MGP_FVBNJ_G0027902	Pnrc1	-0.7998576	0.00105914	0.01999091
MGP_FVBNJ_G0022746	Pros1	-0.7984288	0.00251132	0.03932078
MGP_FVBNJ_G0003760	C130013H08Rik	-0.7917028	0.0031335	0.04638035
MGP_FVBNJ_G0024875	Hectd2	-0.7865744	0.00161649	0.02793272
MGP_FVBNJ_G0017610	Msrb3	-0.7840526	0.00029669	0.0069325
MGP_FVBNJ_G0016266	Mgat5	-0.7811527	0.00122515	0.02255142
MGP_FVBNJ_G0019662	Rhoj	-0.7785862	0.00294195	0.04426075
MGP_FVBNJ_G0031391	Clcn4	-0.7531059	0.00213747	0.03472739
MGP_FVBNJ_G0020768	Pelo	-0.7428328	0.00200595	0.03299976
MGP_FVBNJ_G0020744	Il6st	-0.7347553	0.00025312	0.00605205
MGP_FVBNJ_G0033307	Lonrf1	-0.7338166	0.00070253	0.0144067
MGP_FVBNJ_G0036682	Peg13	-0.7261337	9.33E-05	0.00260433
MGP_FVBNJ_G0021365	Rcbtb2	-0.7208425	0.00078027	0.01571182
MGP_FVBNJ_G0031178	Dusp16	-0.7194257	0.00028639	0.00672703
MGP_FVBNJ_G0033310	AI429214	-0.6832435	0.00162313	0.02802571
MGP_FVBNJ_G0031663	Cyp2a4	-0.6744147	3.87E-05	0.00125202
MGP_FVBNJ_G0033603	Adgre5	-0.6674052	0.00315094	0.04660756
MGP_FVBNJ_G0032154	Akap13	-0.619135	9.38E-05	0.00261488
MGP_FVBNJ_G0031902	Zfp715	-0.5879689	0.00057396	0.01213867
MGP_FVBNJ_G0032301	Ints4	-0.5748215	0.0009581	0.01847499

MGP_FVBNJ_G0021701	A1bg	-0.5172678	3.53E-07	1.85E-05
MGP_FVBNJ_G0016633	Cfap45	-0.4883653	0.00168092	0.02882262
MGP_FVBNJ_G0019709	Susd6	-0.4551219	0.00234125	0.03733105
MGP_FVBNJ_G0024442	NA	-0.4021513	4.92E-20	9.96E-17

Appendix 3: Functional Enrichment of Differentially Expressed Genes Between Nf2-A and $KRAS^{G12D}$-Library Tumours

Functional Enrichment in Downregulated Genes

Gene Ontology: Molecular Function

Source	Term Name	Term ID	P(adj)
GO:MF	cofactor binding	GO:0048037	2.94E-18
GO:MF	heme binding	GO:0020037	1.81E-16
GO:MF	monooxygenase activity	GO:0004497	5.39E-16
GO:MF	tetrapyrrole binding	GO:0046906	6.62E-16
GO:MF	oxidoreductase activity	GO:0016491	3.21E-15
GO:MF	steroid hydroxylase activity	GO:0008395	1.46E-14
GO:MF	iron ion binding	GO:0005506	7.43E-14
GO:MF	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	GO:0016705	2.15E-13
GO:MF	carboxylic acid binding	GO:0031406	5.76E-13
GO:MF	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, reduced flavin or flavoprotein as one donor, and incorporation of one atom of oxygen	GO:0016712	1.43E-12
GO:MF	organic acid binding	GO:0043177	2.17E-12
GO:MF	aromatase activity	GO:0070330	8.81E-12
GO:MF	catalytic activity	GO:0003824	1.21E-10
GO:MF	monocarboxylic acid transmembrane transporter activity	GO:0008028	4.31E-10
GO:MF	lipid transporter activity	GO:0005319	8.98E-10
GO:MF	endopeptidase inhibitor activity	GO:0004866	1.38E-09
GO:MF	vitamin binding	GO:0019842	2.29E-09

GO:MF	monocarboxylic acid binding	GO:0033293	2.50E-09
GO:MF	endopeptidase regulator activity	GO:0061135	2.74E-09
GO:MF	transporter activity	GO:0005215	3.00E-09
GO:MF	peptidase inhibitor activity	GO:0030414	3.29E-09
GO:MF	arachidonic acid monooxygenase activity	GO:0008391	4.47E-09
GO:MF	arachidonic acid epoxygenase activity	GO:0008392	1.23E-08
GO:MF	enzyme inhibitor activity	GO:0004857	1.29E-08
GO:MF	serine-type endopeptidase inhibitor activity	GO:0004867	1.83E-08
GO:MF	peptidase regulator activity	GO:0061134	2.26E-08
GO:MF	lipid binding	GO:0008289	2.32E-08
GO:MF	active transmembrane transporter activity	GO:0022804	2.72E-08
GO:MF	organic anion transmembrane transporter activity	GO:0008514	6.60E-08
GO:MF	secondary active transmembrane transporter activity	GO:0015291	1.17E-07
GO:MF	anion binding	GO:0043168	1.62E-07
GO:MF	symporter activity	GO:0015293	3.13E-07
GO:MF	transmembrane transporter activity	GO:0022857	4.09E-07
GO:MF	serine hydrolase activity	GO:0017171	8.82E-07
GO:MF	hydrolase activity, acting on acid phosphorus-nitrogen bonds	GO:0016825	8.82E-07
GO:MF	anion transmembrane transporter activity	GO:0008509	1.05331E-06
GO:MF	coenzyme binding	GO:0050662	1.18006E-06
GO:MF	serine-type endopeptidase activity	GO:0004252	2.21924E-06
GO:MF	serine-type peptidase activity	GO:0008236	2.717E-06
GO:MF	organic acid transmembrane transporter activity	GO:0005342	2.98101E-06

GO:MF	carboxylic acid transmembrane transporter activity	GO:0046943	2.98101E-06
GO:MF	small molecule binding	GO:0036094	4.61853E-06
GO:MF	glycosaminoglycan binding	GO:0005539	5.11169E-06
GO:MF	protein dimerization activity	GO:0046983	5.8256E-06
GO:MF	heparin binding	GO:0008201	6.16685E-06
GO:MF	sulfur compound binding	GO:1901681	6.6822E-06
GO:MF	transition metal ion binding	GO:0046914	8.83753E-06
GO:MF	oxidoreductase activity, acting on CH-OH group of donors	GO:0016614	1.58614E-05
GO:MF	solute:sodium symporter activity	GO:0015370	2.63422E-05
GO:MF	ion binding	GO:0043167	3.85868E-05
GO:MF	protein homodimerization activity	GO:0042803	4.33958E-05
GO:MF	carboxylic ester hydrolase activity	GO:0052689	4.63241E-05
GO:MF	fatty acid binding	GO:0005504	4.75128E-05
GO:MF	phosphatidylcholine-sterol O-acyltransferase activator activity	GO:0060228	6.60547E-05
GO:MF	bile acid transmembrane transporter activity	GO:0015125	0.000126655
GO:MF	estrogen 16-alpha-hydroxylase activity	GO:0101020	0.000184575
GO:MF	steroid binding	GO:0005496	0.00019423
GO:MF	oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor	GO:0016616	0.000200194
GO:MF	oxidoreductase activity, acting on CH or CH2 groups, quinone or similar compound as acceptor	GO:0033695	0.000269812
GO:MF	caffeine oxidase activity	GO:0034875	0.000269812
GO:MF	triglyceride lipase activity	GO:0004806	0.000311788
GO:MF	signaling receptor binding	GO:0005102	0.000426904

GO:MF	retinol dehydrogenase activity	GO:0004745	0.00045026
GO:MF	amino acid binding	GO:0016597	0.000490893
GO:MF	organic acid:sodium symporter activity	GO:0005343	0.000571179
GO:MF	ion transmembrane transporter activity	GO:0015075	0.000787142
GO:MF	steroid dehydrogenase activity	GO:0016229	0.000788132
GO:MF	pyridoxal phosphate binding	GO:0030170	0.000867427
GO:MF	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, NAD(P)H as one donor, and incorporation of one atom of oxygen	GO:0016709	0.000960714
GO:MF	vitamin B6 binding	GO:0070279	0.001001487
GO:MF	organic hydroxy compound transmembrane transporter activity	GO:1901618	0.001037934
GO:MF	molecular function regulator	GO:0098772	0.0011261
GO:MF	enzyme regulator activity	GO:0030234	0.001303035
GO:MF	identical protein binding	GO:0042802	0.001350589
GO:MF	receptor ligand activity	GO:0048018	0.001595523
GO:MF	protein-lipid complex binding	GO:0071814	0.001746356
GO:MF	lipoprotein particle binding	GO:0071813	0.001746356
GO:MF	lipase inhibitor activity	GO:0055102	0.001997832
GO:MF	solute:cation symporter activity	GO:0015294	0.002027229
GO:MF	signaling receptor activator activity	GO:0030546	0.002061052
GO:MF	retinoic acid 4-hydroxylase activity	GO:0008401	0.002659966
GO:MF	receptor regulator activity	GO:0030545	0.002665111
GO:MF	quaternary ammonium group binding	GO:0050997	0.003123814

GO:MF	phosphatidylcholine binding	GO:0031210	0.003123814
GO:MF	alcohol binding	GO:0043178	0.003229953
GO:MF	lipoprotein particle receptor binding	GO:0070325	0.003297245
GO:MF	transforming growth factor beta receptor binding	GO:0005160	0.003401295
GO:MF	growth factor activity	GO:0008083	0.003411676
GO:MF	extracellular matrix structural constituent	GO:0005201	0.003736731
GO:MF	oxidoreductase activity, acting on CH or CH2 groups	GO:0016725	0.00380783
GO:MF	sterol esterase activity	GO:0004771	0.003869006
GO:MF	transferase activity, transferring sulfur-containing groups	GO:0016782	0.004290363
GO:MF	calcium ion binding	GO:0005509	0.004642747
GO:MF	complement binding	GO:0001848	0.004709524
GO:MF	nerve growth factor binding	GO:0048406	0.005153332
GO:MF	sodium ion transmembrane transporter activity	GO:0015081	0.006326919
GO:MF	phospholipid binding	GO:0005543	0.006579004
GO:MF	fatty acid ligase activity	GO:0015645	0.006762123
GO:MF	flavin adenine dinucleotide binding	GO:0050660	0.007676817
GO:MF	active ion transmembrane transporter activity	GO:0022853	0.008050288
GO:MF	brain-derived neurotrophic factor binding	GO:0048403	0.008120362
GO:MF	carbohydrate binding	GO:0030246	0.009190421
GO:MF	glucuronosyltransferase activity	GO:0015020	0.009345906
GO:MF	oxidoreductase activity, acting on single donors with incorporation of molecular oxygen, incorporation of two atoms of oxygen	GO:0016702	0.00960913
GO:MF	oxidoreductase activity, acting on single donors with incorporation of molecular oxygen	GO:0016701	0.012620029

GO:MF	uracil binding	GO:0002058	0.012710687
GO:MF	pyrimidine nucleobase binding	GO:0002061	0.012710687
GO:MF	sulfotransferase activity	GO:0008146	0.013970453
GO:MF	transferase activity, transferring acyl groups	GO:0016746	0.014398355
GO:MF	bile acid binding	GO:0032052	0.015598446
GO:MF	long-chain fatty acid transporter activity	GO:0005324	0.016325425
GO:MF	CoA-ligase activity	GO:0016405	0.017579955
GO:MF	peptide hormone binding	GO:0017046	0.018523039
GO:MF	neurotrophin binding	GO:0043121	0.019905723
GO:MF	fatty-acyl-CoA synthase activity	GO:0004321	0.02035698
GO:MF	acyl-CoA ligase activity	GO:0003996	0.021262305
GO:MF	chemoattractant activity	GO:0042056	0.02218264
GO:MF	steroid dehydrogenase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor	GO:0033764	0.022542672
GO:MF	cell adhesion molecule binding	GO:0050839	0.023313447
GO:MF	alditol:NADP+ 1-oxidoreductase activity	GO:0004032	0.023947867
GO:MF	alanine-glyoxylate transaminase activity	GO:0008453	0.026750493
GO:MF	monosaccharide binding	GO:0048029	0.027041252
GO:MF	ketosteroid monooxygenase activity	GO:0047086	0.028855505
GO:MF	lyase activity	GO:0016829	0.030565167
GO:MF	inorganic molecular entity transmembrane transporter activity	GO:0015318	0.033052529
GO:MF	glycine N-acyltransferase activity	GO:0047961	0.033327152
GO:MF	vitamin E binding	GO:0008431	0.033734534
GO:MF	cation binding	GO:0043169	0.037097072
GO:MF	acid-thiol ligase activity	GO:0016878	0.038551003
GO:MF	endopeptidase activity	GO:0004175	0.038725896
GO:MF	high-density lipoprotein particle receptor binding	GO:0070653	0.038979289

GO:MF	hormone binding	GO:0042562	0.045020138
GO:MF	amide binding	GO:0033218	0.046073638

Gene Ontology: Biological Process			
GO:BP	organic acid metabolic process	GO:0006082	2.10E-46
GO:BP	carboxylic acid metabolic process	GO:0019752	6.84E-44
GO:BP	oxoacid metabolic process	GO:0043436	3.42E-42
GO:BP	small molecule metabolic process	GO:0044281	4.90E-37
GO:BP	monocarboxylic acid metabolic process	GO:0032787	1.25E-31
GO:BP	lipid metabolic process	GO:0006629	7.09E-27
GO:BP	drug metabolic process	GO:0017144	3.33E-24
GO:BP	fatty acid metabolic process	GO:0006631	4.19E-20
GO:BP	drug catabolic process	GO:0042737	8.42E-20
GO:BP	organic hydroxy compound metabolic process	GO:1901615	2.60E-19
GO:BP	cellular lipid metabolic process	GO:0044255	1.19E-18
GO:BP	small molecule catabolic process	GO:0044282	2.10E-18
GO:BP	steroid metabolic process	GO:0008202	1.79E-17
GO:BP	oxidation-reduction process	GO:0055114	9.71E-17
GO:BP	fibrinolysis	GO:0042730	1.37E-16
GO:BP	blood coagulation	GO:0007596	2.75E-16
GO:BP	hemostasis	GO:0007599	3.44E-16
GO:BP	coagulation	GO:0050817	3.85E-16
GO:BP	alpha-amino acid metabolic process	GO:1901605	2.15E-15
GO:BP	organic acid catabolic process	GO:0016054	5.06E-15
GO:BP	carboxylic acid catabolic process	GO:0046395	5.06E-15
GO:BP	alpha-amino acid catabolic process	GO:1901606	6.58E-15
GO:BP	negative regulation of blood coagulation	GO:0030195	1.77E-14
GO:BP	regulation of blood coagulation	GO:0030193	2.29E-14

GO:BP	negative regulation of hemostasis	GO:1900047	2.50E-14
GO:BP	organic acid biosynthetic process	GO:0016053	2.60E-14
GO:BP	regulation of hemostasis	GO:1900046	2.74E-14
GO:BP	negative regulation of coagulation	GO:0050819	3.50E-14
GO:BP	cellular amino acid metabolic process	GO:0006520	3.79E-14
GO:BP	regulation of coagulation	GO:0050818	3.89E-14
GO:BP	regulation of biological quality	GO:0065008	5.35E-14
GO:BP	cellular amino acid catabolic process	GO:0009063	8.43E-14
GO:BP	lipid localization	GO:0010876	1.17E-13
GO:BP	carboxylic acid biosynthetic process	GO:0046394	1.74E-13
GO:BP	triglyceride metabolic process	GO:0006641	5.54E-13
GO:BP	small molecule biosynthetic process	GO:0044283	5.70E-13
GO:BP	exogenous drug catabolic process	GO:0042738	6.64E-13
GO:BP	cellular ketone metabolic process	GO:0042180	9.43E-13
GO:BP	organic hydroxy compound transport	GO:0015850	1.09E-12
GO:BP	organic anion transport	GO:0015711	1.32E-12
GO:BP	lipid transport	GO:0006869	2.64E-12
GO:BP	xenobiotic metabolic process	GO:0006805	3.29E-12
GO:BP	aromatic amino acid family catabolic process	GO:0009074	3.56E-12
GO:BP	acylglycerol metabolic process	GO:0006639	4.99E-12
GO:BP	negative regulation of wound healing	GO:0061045	7.54E-12
GO:BP	neutral lipid metabolic process	GO:0006638	8.21E-12
GO:BP	cofactor metabolic process	GO:0051186	1.12E-11
GO:BP	wound healing	GO:0042060	1.76E-11

GO:BP	regulation of body fluid levels	GO:0050878	1.83E-11
GO:BP	catabolic process	GO:0009056	2.64E-11
GO:BP	response to external stimulus	GO:0009605	3.21E-11
GO:BP	alcohol metabolic process	GO:0006066	4.53E-11
GO:BP	plasminogen activation	GO:0031639	4.56E-11
GO:BP	regulation of hormone levels	GO:0010817	5.49E-11
GO:BP	cellular response to xenobiotic stimulus	GO:0071466	7.63E-11
GO:BP	response to xenobiotic stimulus	GO:0009410	8.87E-11
GO:BP	cellular catabolic process	GO:0044248	1.07E-10
GO:BP	negative regulation of response to wounding	GO:1903035	1.22E-10
GO:BP	regulation of wound healing	GO:0061041	1.25E-10
GO:BP	positive regulation of lipid catabolic process	GO:0050996	1.98E-10
GO:BP	long-chain fatty acid metabolic process	GO:0001676	4.11E-10
GO:BP	aromatic amino acid family metabolic process	GO:0009072	4.40E-10
GO:BP	regulation of fibrinolysis	GO:0051917	4.50E-10
GO:BP	arachidonic acid metabolic process	GO:0019369	8.12E-10
GO:BP	coenzyme metabolic process	GO:0006732	8.67E-10
GO:BP	cellular response to chemical stimulus	GO:0070887	8.98E-10
GO:BP	anion transport	GO:0006820	1.15E-09
GO:BP	ion transport	GO:0006811	1.44E-09
GO:BP	positive regulation of lipid metabolic process	GO:0045834	1.51E-09
GO:BP	response to wounding	GO:0009611	1.88E-09
GO:BP	unsaturated fatty acid metabolic process	GO:0033559	2.10E-09
GO:BP	regulation of response to wounding	GO:1903034	2.16E-09
GO:BP	response to drug	GO:0042493	3.77E-09
GO:BP	purine-containing compound metabolic process	GO:0072521	4.38E-09

GO:BP	zymogen activation	GO:0031638	4.44E-09
GO:BP	lipid homeostasis	GO:0055088	4.81E-09
GO:BP	acute-phase response	GO:0006953	4.94E-09
GO:BP	regulation of lipid metabolic process	GO:0019216	5.02E-09
GO:BP	lipid biosynthetic process	GO:0008610	5.10E-09
GO:BP	hormone metabolic process	GO:0042445	6.40E-09
GO:BP	regulation of lipid catabolic process	GO:0050994	6.50E-09
GO:BP	monocarboxylic acid biosynthetic process	GO:0072330	7.73E-09
GO:BP	cholesterol metabolic process	GO:0008203	8.41E-09
GO:BP	lipid catabolic process	GO:0016042	1.30E-08
GO:BP	negative regulation of proteolysis	GO:0045861	1.62E-08
GO:BP	icosanoid metabolic process	GO:0006690	1.72E-08
GO:BP	negative regulation of endopeptidase activity	GO:0010951	1.83E-08
GO:BP	epoxygenase P450 pathway	GO:0019373	1.98E-08
GO:BP	negative regulation of peptidase activity	GO:0010466	4.38E-08
GO:BP	acylglycerol homeostasis	GO:0055090	4.71E-08
GO:BP	triglyceride homeostasis	GO:0070328	4.71E-08
GO:BP	fatty acid derivative metabolic process	GO:1901568	5.57E-08
GO:BP	nucleobase-containing small molecule metabolic process	GO:0055086	5.92E-08
GO:BP	steroid biosynthetic process	GO:0006694	7.33E-08
GO:BP	chemical homeostasis	GO:0048878	9.01E-08
GO:BP	sterol metabolic process	GO:0016125	1.07E-07
GO:BP	secondary alcohol metabolic process	GO:1902652	1.07E-07
GO:BP	sulfur compound metabolic process	GO:0006790	1.15E-07
GO:BP	defense response	GO:0006952	1.23E-07
GO:BP	negative regulation of hydrolase activity	GO:0051346	1.37E-07
GO:BP	negative regulation of response to external stimulus	GO:0032102	1.37E-07

GO:BP	blood coagulation, fibrin clot formation	GO:0072378	1.64E-07
GO:BP	regulation of cellular ketone metabolic process	GO:0010565	1.86E-07
GO:BP	acute inflammatory response	GO:0002526	1.86E-07
GO:BP	regulation of lipid biosynthetic process	GO:0046890	2.30E-07
GO:BP	organic substance catabolic process	GO:1901575	2.53E-07
GO:BP	triglyceride catabolic process	GO:0019433	3.44E-07
GO:BP	neutral lipid catabolic process	GO:0046461	3.69E-07
GO:BP	acylglycerol catabolic process	GO:0046464	3.69E-07
GO:BP	plasma lipoprotein particle remodeling	GO:0034369	4.84E-07
GO:BP	protein-lipid complex remodeling	GO:0034368	4.84E-07
GO:BP	regulation of triglyceride metabolic process	GO:0090207	5.44E-07
GO:BP	protein activation cascade	GO:0072376	7.88E-07
GO:BP	protein-containing complex remodeling	GO:0034367	9.73E-07
GO:BP	ammonium ion metabolic process	GO:0097164	9.81E-07
GO:BP	organic acid transport	GO:0015849	9.91E-07
GO:BP	glycerolipid metabolic process	GO:0046486	1.11035E-06
GO:BP	phosphatidylcholine metabolic process	GO:0046470	1.11458E-06
GO:BP	very-low-density lipoprotein particle remodeling	GO:0034372	1.11852E-06
GO:BP	plasma lipoprotein particle organization	GO:0071827	1.34157E-06
GO:BP	regulation of multicellular organismal process	GO:0051239	1.34631E-06
GO:BP	positive regulation of lipid biosynthetic process	GO:0046889	1.46307E-06

GO:BP	regulation of steroid metabolic process	GO:0019218	2.39558E-06
GO:BP	negative regulation of fibrinolysis	GO:0051918	2.39951E-06
GO:BP	glycerolipid catabolic process	GO:0046503	2.48121E-06
GO:BP	carboxylic acid transport	GO:0046942	2.54306E-06
GO:BP	triglyceride-rich lipoprotein particle remodeling	GO:0034370	3.30946E-06
GO:BP	monocarboxylic acid transport	GO:0015718	3.34284E-06
GO:BP	regulation of triglyceride catabolic process	GO:0010896	3.45376E-06
GO:BP	cellular lipid catabolic process	GO:0044242	3.87653E-06
GO:BP	humoral immune response	GO:0006959	4.02063E-06
GO:BP	organic hydroxy compound biosynthetic process	GO:1901617	4.18209E-06
GO:BP	positive regulation of cholesterol esterification	GO:0010873	4.19864E-06
GO:BP	positive regulation of transport	GO:0051050	4.20957E-06
GO:BP	regulation of small molecule metabolic process	GO:0062012	4.3995E-06
GO:BP	cellular hormone metabolic process	GO:0034754	4.61963E-06
GO:BP	sterol esterification	GO:0034434	4.98705E-06
GO:BP	cholesterol esterification	GO:0034435	4.98705E-06
GO:BP	steroid esterification	GO:0034433	4.98705E-06
GO:BP	primary alcohol metabolic process	GO:0034308	5.02236E-06
GO:BP	cellular modified amino acid metabolic process	GO:0006575	5.22235E-06
GO:BP	response to metal ion	GO:0010038	6.35326E-06
GO:BP	thioester metabolic process	GO:0035383	6.42013E-06
GO:BP	acyl-CoA metabolic process	GO:0006637	6.42013E-06
GO:BP	positive regulation of triglyceride metabolic process	GO:0090208	6.49586E-06
GO:BP	regulation of endopeptidase activity	GO:0052548	6.55977E-06

GO:BP	proteolysis	GO:0006508	7.05622E-06
GO:BP	protein-lipid complex subunit organization	GO:0071825	7.51614E-06
GO:BP	response to inorganic substance	GO:0010035	7.82436E-06
GO:BP	high-density lipoprotein particle assembly	GO:0034380	8.16001E-06
GO:BP	regulation of lipase activity	GO:0060191	8.81996E-06
GO:BP	regulation of neurotransmitter levels	GO:0001505	1.09166E-05
GO:BP	bile acid and bile salt transport	GO:0015721	1.17395E-05
GO:BP	negative regulation of catalytic activity	GO:0043086	1.3588E-05
GO:BP	antibiotic metabolic process	GO:0016999	1.49519E-05
GO:BP	response to glucagon	GO:0033762	1.50063E-05
GO:BP	regulation of peptidase activity	GO:0052547	1.51656E-05
GO:BP	protein processing	GO:0016485	1.71534E-05
GO:BP	protein maturation	GO:0051604	1.72423E-05
GO:BP	monosaccharide metabolic process	GO:0005996	1.74292E-05
GO:BP	positive regulation of steroid metabolic process	GO:0045940	1.76246E-05
GO:BP	positive regulation of triglyceride catabolic process	GO:0010898	1.77055E-05
GO:BP	positive regulation of heterotypic cell-cell adhesion	GO:0034116	1.81893E-05
GO:BP	urate metabolic process	GO:0046415	1.93978E-05
GO:BP	positive regulation of fibrinolysis	GO:0051919	1.96562E-05
GO:BP	regulation of localization	GO:0032879	2.18579E-05
GO:BP	regulation of lipoprotein lipase activity	GO:0051004	2.89854E-05
GO:BP	regulation of plasma lipoprotein particle levels	GO:0097006	2.90203E-05
GO:BP	response to organic substance	GO:0010033	3.04325E-05

GO:BP	nucleoside bisphosphate metabolic process	GO:0033865	3.14154E-05
GO:BP	purine nucleoside bisphosphate metabolic process	GO:0034032	3.14154E-05
GO:BP	ribonucleoside bisphosphate metabolic process	GO:0033875	3.14154E-05
GO:BP	cholesterol homeostasis	GO:0042632	3.19801E-05
GO:BP	cell adhesion	GO:0007155	3.31648E-05
GO:BP	bile acid metabolic process	GO:0008206	3.552E-05
GO:BP	liver development	GO:0001889	3.71224E-05
GO:BP	blood circulation	GO:0008015	3.77545E-05
GO:BP	sterol homeostasis	GO:0055092	3.84633E-05
GO:BP	cholesterol transport	GO:0030301	3.87561E-05
GO:BP	system development	GO:0048731	4.05585E-05
GO:BP	innate immune response	GO:0045087	4.11117E-05
GO:BP	regulation of cholesterol esterification	GO:0010872	4.42323E-05
GO:BP	monosaccharide biosynthetic process	GO:0046364	4.67513E-05
GO:BP	hepaticobiliary system development	GO:0061008	4.84173E-05
GO:BP	benzene-containing compound metabolic process	GO:0042537	5.09469E-05
GO:BP	response to other organism	GO:0051707	5.22829E-05
GO:BP	biological adhesion	GO:0022610	5.32217E-05
GO:BP	sterol transport	GO:0015918	5.49365E-05
GO:BP	plasma lipoprotein particle assembly	GO:0034377	5.628E-05
GO:BP	organophosphate metabolic process	GO:0019637	5.77856E-05
GO:BP	response to external biotic stimulus	GO:0043207	5.88408E-05
GO:BP	inflammatory response	GO:0006954	6.10617E-05
GO:BP	secretion	GO:0046903	6.49589E-05
GO:BP	nucleoside phosphate metabolic process	GO:0006753	6.79161E-05
GO:BP	response to stress	GO:0006950	7.06252E-05

GO:BP	complement activation, alternative pathway	GO:0006957	7.32701E-05
GO:BP	response to biotic stimulus	GO:0009607	8.06427E-05
GO:BP	circulatory system process	GO:0003013	8.19458E-05
GO:BP	adaptive immune response	GO:0002250	8.60495E-05
GO:BP	bile acid biosynthetic process	GO:0006699	8.85134E-05
GO:BP	cellular response to glucagon stimulus	GO:0071377	9.06236E-05
GO:BP	organonitrogen compound metabolic process	GO:1901564	9.73302E-05
GO:BP	positive regulation of response to stimulus	GO:0048584	9.94951E-05
GO:BP	negative regulation of protein metabolic process	GO:0051248	9.95299E-05
GO:BP	terpenoid metabolic process	GO:0006721	0.000108737
GO:BP	positive regulation of small molecule metabolic process	GO:0062013	0.000112219
GO:BP	kynurenine metabolic process	GO:0070189	0.000112498
GO:BP	nucleotide metabolic process	GO:0009117	0.000122615
GO:BP	regulation of heterotypic cell-cell adhesion	GO:0034114	0.000129703
GO:BP	negative regulation of extrinsic apoptotic signaling pathway via death domain receptors	GO:1902042	0.000145895
GO:BP	negative regulation of endothelial cell apoptotic process	GO:2000352	0.000148622
GO:BP	positive regulation of fatty acid metabolic process	GO:0045923	0.000149433
GO:BP	positive regulation of fatty acid biosynthetic process	GO:0045723	0.000153946
GO:BP	regulation of lipid localization	GO:1905952	0.000156635
GO:BP	transmembrane transport	GO:0055085	0.000168604
GO:BP	positive regulation of hemostasis	GO:1900048	0.000171427

GO:BP	positive regulation of blood coagulation	GO:0030194	0.000171427
GO:BP	metal ion transport	GO:0030001	0.000187967
GO:BP	complement activation	GO:0006956	0.000205522
GO:BP	regulation of response to external stimulus	GO:0032101	0.000229885
GO:BP	defense response to other organism	GO:0098542	0.000246088
GO:BP	high-density lipoprotein particle remodeling	GO:0034375	0.000253594
GO:BP	positive regulation of coagulation	GO:0050820	0.000254418
GO:BP	regulation of fatty acid biosynthetic process	GO:0042304	0.000272666
GO:BP	response to hormone	GO:0009725	0.000334323
GO:BP	regulation of protein metabolic process	GO:0051246	0.000337463
GO:BP	regulation of proteolysis	GO:0030162	0.000345176
GO:BP	negative regulation of cellular protein metabolic process	GO:0032269	0.000363456
GO:BP	purine nucleotide metabolic process	GO:0006163	0.000377568
GO:BP	regulation of blood vessel diameter	GO:0097746	0.00038488
GO:BP	regulation of tube diameter	GO:0035296	0.00038488
GO:BP	animal organ development	GO:0048513	0.00041361
GO:BP	protein-lipid complex assembly	GO:0065005	0.000427651
GO:BP	regulation of tube size	GO:0035150	0.000438147
GO:BP	anion homeostasis	GO:0055081	0.000442449
GO:BP	positive regulation of lipase activity	GO:0060193	0.000448665
GO:BP	organophosphate ester transport	GO:0015748	0.00045832
GO:BP	response to oxygen-containing compound	GO:1901700	0.000475071
GO:BP	extracellular structure organization	GO:0043062	0.000485741
GO:BP	neurotransmitter metabolic process	GO:0042133	0.000508551
GO:BP	regulation of cellular protein metabolic process	GO:0032268	0.000527465

GO:BP	negative regulation of molecular function	GO:0044092	0.000529309
GO:BP	tube development	GO:0035295	0.000550239
GO:BP	regulation of transport	GO:0051049	0.000573562
GO:BP	tyrosine metabolic process	GO:0006570	0.000592044
GO:BP	positive regulation of vasoconstriction	GO:0045907	0.00061416
GO:BP	tryptophan catabolic process	GO:0006569	0.000678605
GO:BP	indolalkylamine catabolic process	GO:0046218	0.000678605
GO:BP	indole-containing compound catabolic process	GO:0042436	0.000678605
GO:BP	purine ribonucleotide metabolic process	GO:0009150	0.000681418
GO:BP	regulation of extrinsic apoptotic signaling pathway via death domain receptors	GO:1902041	0.000705977
GO:BP	negative regulation of epithelial cell apoptotic process	GO:1904036	0.000755124
GO:BP	regulation of fatty acid metabolic process	GO:0019217	0.00075856
GO:BP	diterpenoid metabolic process	GO:0016101	0.000774544
GO:BP	regulation of complement activation	GO:0030449	0.000797811
GO:BP	platelet aggregation	GO:0070527	0.000806495
GO:BP	positive regulation of multicellular organismal process	GO:0051240	0.000881711
GO:BP	serine family amino acid metabolic process	GO:0009069	0.000953272
GO:BP	complement activation, lectin pathway	GO:0001867	0.000971476
GO:BP	cellular amino acid biosynthetic process	GO:0008652	0.001013437
GO:BP	antibiotic catabolic process	GO:0017001	0.001025204
GO:BP	alcohol catabolic process	GO:0046164	0.001099993
GO:BP	phospholipid efflux	GO:0033700	0.001121155

GO:BP	ion homeostasis	GO:0050801	0.001164272
GO:BP	isoprenoid metabolic process	GO:0006720	0.001171341
GO:BP	cellular biogenic amine catabolic process	GO:0042402	0.001201566
GO:BP	response to endogenous stimulus	GO:0009719	0.001205194
GO:BP	regulation of endothelial cell apoptotic process	GO:2000351	0.001232469
GO:BP	induction of bacterial agglutination	GO:0043152	0.001273132
GO:BP	multicellular organism development	GO:0007275	0.001310518
GO:BP	glucose metabolic process	GO:0006006	0.001319637
GO:BP	retinoid metabolic process	GO:0001523	0.001348509
GO:BP	vascular process in circulatory system	GO:0003018	0.00141079
GO:BP	heterotypic cell-cell adhesion	GO:0034113	0.001453262
GO:BP	ribonucleotide metabolic process	GO:0009259	0.001502173
GO:BP	negative regulation of multicellular organismal process	GO:0051241	0.001528684
GO:BP	endothelial cell apoptotic process	GO:0072577	0.001532314
GO:BP	positive regulation of cell communication	GO:0010647	0.001556855
GO:BP	gluconeogenesis	GO:0006094	0.001556963
GO:BP	organic cyclic compound catabolic process	GO:1901361	0.001676385
GO:BP	cation transport	GO:0006812	0.001718658
GO:BP	indole-containing compound metabolic process	GO:0042430	0.001839317
GO:BP	regulation of ion transport	GO:0043269	0.001852735
GO:BP	glutamine family amino acid metabolic process	GO:0009064	0.001913991
GO:BP	phospholipid transport	GO:0015914	0.001914125
GO:BP	positive regulation of signaling	GO:0023056	0.00193335

GO:BP	response to extracellular stimulus	GO:0009991	0.002040245
GO:BP	hexose metabolic process	GO:0019318	0.002083928
GO:BP	hexose biosynthetic process	GO:0019319	0.002156866
GO:BP	animal organ regeneration	GO:0031100	0.002305934
GO:BP	thyroid hormone transport	GO:0070327	0.00231767
GO:BP	dicarboxylic acid metabolic process	GO:0043648	0.002361415
GO:BP	homotypic cell-cell adhesion	GO:0034109	0.002375937
GO:BP	organic substance transport	GO:0071702	0.002377613
GO:BP	C21-steroid hormone metabolic process	GO:0008207	0.00250696
GO:BP	ribose phosphate metabolic process	GO:0019693	0.002530683
GO:BP	positive regulation of transmembrane receptor protein serine/threonine kinase signaling pathway	GO:0090100	0.002603301
GO:BP	humoral immune response mediated by circulating immunoglobulin	GO:0002455	0.002672684
GO:BP	amine catabolic process	GO:0009310	0.002859459
GO:BP	regulation of vasoconstriction	GO:0019229	0.00295586
GO:BP	extrinsic apoptotic signaling pathway via death domain receptors	GO:0008625	0.003211877
GO:BP	erythrose 4-phosphate/phosphoenolpyruvate family amino acid catabolic process	GO:1902222	0.003267641
GO:BP	L-phenylalanine catabolic process	GO:0006559	0.003267641
GO:BP	vitamin transport	GO:0051180	0.003384496
GO:BP	response to acid chemical	GO:0001101	0.003403398
GO:BP	positive regulation of phosphate metabolic process	GO:0045937	0.003419052

GO:BP	positive regulation of phosphorus metabolic process	GO:0010562	0.003419052
GO:BP	anatomical structure development	GO:0048856	0.003446674
GO:BP	regulation of very-low-density lipoprotein particle remodeling	GO:0010901	0.003591326
GO:BP	negative regulation of lipase activity	GO:0060192	0.003721554
GO:BP	localization	GO:0051179	0.003878052
GO:BP	fatty acid biosynthetic process	GO:0006633	0.003922208
GO:BP	calcium ion-regulated exocytosis of neurotransmitter	GO:0048791	0.004006762
GO:BP	cellular chemical homeostasis	GO:0055082	0.004086435
GO:BP	platelet activation	GO:0030168	0.004223212
GO:BP	immune response	GO:0006955	0.004225801
GO:BP	monocarboxylic acid catabolic process	GO:0072329	0.004499481
GO:BP	response to lipid	GO:0033993	0.004513293
GO:BP	nucleobase metabolic process	GO:0009112	0.005025523
GO:BP	drug transport	GO:0015893	0.005077318
GO:BP	regulation of lipid transport	GO:0032368	0.005187058
GO:BP	regulation of epithelial cell apoptotic process	GO:1904035	0.005242131
GO:BP	response to lipopolysaccharide	GO:0032496	0.005254181
GO:BP	histidine catabolic process to glutamate and formate	GO:0019557	0.005342756
GO:BP	formamide metabolic process	GO:0043606	0.005342756
GO:BP	histidine catabolic process to glutamate and formamide	GO:0019556	0.005342756
GO:BP	response to steroid hormone	GO:0048545	0.005432806
GO:BP	retinol metabolic process	GO:0042572	0.005485616
GO:BP	cytolysis	GO:0019835	0.005509245

GO:BP	cholesterol efflux	GO:0033344	0.005695002
GO:BP	regulation of blood circulation	GO:1903522	0.005768126
GO:BP	positive regulation of blood circulation	GO:1903524	0.006004803
GO:BP	alpha-amino acid biosynthetic process	GO:1901607	0.006054047
GO:BP	response to corticosteroid	GO:0031960	0.006146522
GO:BP	aspartate family amino acid metabolic process	GO:0009066	0.006196265
GO:BP	negative regulation of extrinsic apoptotic signaling pathway	GO:2001237	0.006206332
GO:BP	positive regulation of phosphorylation	GO:0042327	0.006257498
GO:BP	drug transmembrane transport	GO:0006855	0.006327604
GO:BP	L-phenylalanine metabolic process	GO:0006558	0.006460914
GO:BP	erythrose 4-phosphate/phosphoenolpyruvate family amino acid metabolic process	GO:1902221	0.006460914
GO:BP	vasoconstriction	GO:0042310	0.006622755
GO:BP	positive regulation of ion transport	GO:0043270	0.006640075
GO:BP	response to nutrient levels	GO:0031667	0.006675636
GO:BP	phosphate-containing compound metabolic process	GO:0006796	0.006798671
GO:BP	hyaluronan metabolic process	GO:0030212	0.006835205
GO:BP	complement activation, classical pathway	GO:0006958	0.006836306
GO:BP	positive regulation of lipoprotein lipase activity	GO:0051006	0.007052521
GO:BP	positive regulation of exocytosis	GO:0045921	0.007057316
GO:BP	cellular response to metal ion	GO:0071248	0.007107685

GO:BP	aromatic compound catabolic process	GO:0019439	0.007322062
GO:BP	carbohydrate metabolic process	GO:0005975	0.00736055
GO:BP	phenol-containing compound metabolic process	GO:0018958	0.007520705
GO:BP	reverse cholesterol transport	GO:0043691	0.007721618
GO:BP	NAD biosynthetic process	GO:0009435	0.007774349
GO:BP	plasma lipoprotein particle clearance	GO:0034381	0.007987242
GO:BP	response to glucocorticoid	GO:0051384	0.008281613
GO:BP	'de novo' NAD biosynthetic process	GO:0034627	0.008313926
GO:BP	'de novo' NAD biosynthetic process from tryptophan	GO:0034354	0.008313926
GO:BP	response to peptide hormone	GO:0043434	0.008381551
GO:BP	negative regulation of blood vessel diameter	GO:0097756	0.008473614
GO:BP	response to chemical	GO:0042221	0.008484661
GO:BP	lipid modification	GO:0030258	0.008710205
GO:BP	detoxification	GO:0098754	0.008760818
GO:BP	indolalkylamine metabolic process	GO:0006586	0.00890632
GO:BP	tryptophan metabolic process	GO:0006568	0.00890632
GO:BP	response to bacterium	GO:0009617	0.009034322
GO:BP	organonitrogen compound catabolic process	GO:1901565	0.009209163
GO:BP	vitamin metabolic process	GO:0006766	0.00927403
GO:BP	transport	GO:0006810	0.009311716
GO:BP	cellular aldehyde metabolic process	GO:0006081	0.00943343
GO:BP	regulation of humoral immune response	GO:0002920	0.00957931
GO:BP	organic hydroxy compound catabolic process	GO:1901616	0.009733872
GO:BP	regulation of hydrolase activity	GO:0051336	0.009744944

GO:BP	homeostatic process	GO:0042592	0.009811179
GO:BP	intestinal lipid absorption	GO:0098856	0.010031171
GO:BP	regulation of intestinal cholesterol absorption	GO:0030300	0.010520931
GO:BP	positive regulation of pathway-restricted SMAD protein phosphorylation	GO:0010862	0.011037523
GO:BP	phosphorus metabolic process	GO:0006793	0.011227765
GO:BP	regulation of secretion	GO:0051046	0.011565454
GO:BP	endocrine pancreas development	GO:0031018	0.011659509
GO:BP	epithelial cell apoptotic process	GO:1904019	0.011847014
GO:BP	amine transport	GO:0015837	0.012484108
GO:BP	positive regulation of wound healing	GO:0090303	0.01270545
GO:BP	developmental process	GO:0032502	0.012744989
GO:BP	second-messenger-mediated signaling	GO:0019932	0.012868054
GO:BP	glutamate metabolic process	GO:0006536	0.013146555
GO:BP	response to molecule of bacterial origin	GO:0002237	0.013194376
GO:BP	histidine catabolic process	GO:0006548	0.013273771
GO:BP	formate metabolic process	GO:0015942	0.013273771
GO:BP	positive regulation of triglyceride lipase activity	GO:0061365	0.013566605
GO:BP	cellular response to inorganic substance	GO:0071241	0.013634911
GO:BP	positive regulation of peptide hormone secretion	GO:0090277	0.013824451
GO:BP	cholesterol biosynthetic process	GO:0006695	0.014210206
GO:BP	response to organic cyclic compound	GO:0014070	0.01421776
GO:BP	L-lysine catabolic process	GO:0019477	0.015120579
GO:BP	lysine catabolic process	GO:0006554	0.015120579
GO:BP	L-lysine metabolic process	GO:0046440	0.015120579
GO:BP	L-lysine catabolic process to acetyl-CoA	GO:0019474	0.015120579

GO:BP	neurotransmitter transport	GO:0006836	0.015343174
GO:BP	sulfur amino acid metabolic process	GO:0000096	0.016262145
GO:BP	canalicular bile acid transport	GO:0015722	0.017234219
GO:BP	aspartate family amino acid catabolic process	GO:0009068	0.017238471
GO:BP	export from cell	GO:0140352	0.017385224
GO:BP	tissue development	GO:0009888	0.017520659
GO:BP	negative regulation of lipid metabolic process	GO:0045833	0.017724289
GO:BP	regulation of response to stimulus	GO:0048583	0.0177968
GO:BP	L-kynurenine metabolic process	GO:0097052	0.018275104
GO:BP	epithelium development	GO:0060429	0.018545623
GO:BP	regulation of intestinal lipid absorption	GO:1904729	0.018693804
GO:BP	regulation of secretion by cell	GO:1903530	0.018757552
GO:BP	regulation of sterol transport	GO:0032371	0.019950307
GO:BP	regulation of cholesterol transport	GO:0032374	0.019950307
GO:BP	locomotion	GO:0040011	0.020104939
GO:BP	pyruvate metabolic process	GO:0006090	0.020614871
GO:BP	chylomicron remnant clearance	GO:0034382	0.020713492
GO:BP	triglyceride-rich lipoprotein particle clearance	GO:0071830	0.020713492
GO:BP	regulation of cellular component movement	GO:0051270	0.020865025
GO:BP	defense response to fungus	GO:0050832	0.023009525
GO:BP	negative regulation of small molecule metabolic process	GO:0062014	0.023479906
GO:BP	fatty acid catabolic process	GO:0009062	0.023701329
GO:BP	negative regulation of humoral immune response	GO:0002921	0.025439936
GO:BP	anthranilate metabolic process	GO:0043420	0.025703366

GO:BP	imidazole-containing compound catabolic process	GO:0052805	0.026382402
GO:BP	regulation of plasminogen activation	GO:0010755	0.026445609
GO:BP	inorganic ion homeostasis	GO:0098771	0.026735005
GO:BP	secondary alcohol biosynthetic process	GO:1902653	0.026898346
GO:BP	acetyl-CoA metabolic process	GO:0006084	0.027688718
GO:BP	regulation of extrinsic apoptotic signaling pathway	GO:2001236	0.02867571
GO:BP	sodium ion transport	GO:0006814	0.028804206
GO:BP	response to calcium ion	GO:0051592	0.02924127
GO:BP	regulation of extracellular matrix constituent secretion	GO:0003330	0.02924127
GO:BP	quinolinate metabolic process	GO:0046874	0.029952568
GO:BP	positive regulation of ERK1 and ERK2 cascade	GO:0070374	0.030059142
GO:BP	regeneration	GO:0031099	0.031148813
GO:BP	positive regulation of hormone secretion	GO:0046887	0.031577299
GO:BP	immunoglobulin mediated immune response	GO:0016064	0.031902634
GO:BP	regulation of response to stress	GO:0080134	0.032435475
GO:BP	cell-cell adhesion	GO:0098609	0.032683374
GO:BP	hormone transport	GO:0009914	0.033516364
GO:BP	B cell mediated immunity	GO:0019724	0.036529112
GO:BP	cell killing	GO:0001906	0.037404636
GO:BP	lysine metabolic process	GO:0006553	0.037467368
GO:BP	immune system process	GO:0002376	0.038758689
GO:BP	cellular response to oxygen-containing compound	GO:1901701	0.040552312
GO:BP	positive regulation of response to wounding	GO:1903036	0.043430956
GO:BP	carbohydrate biosynthetic process	GO:0016051	0.043600137
GO:BP	positive regulation of protein phosphorylation	GO:0001934	0.044390565

GO:BP	fatty acid homeostasis	GO:0055089	0.044395646
GO:BP	cellular homeostasis	GO:0019725	0.04442799
GO:BP	xenobiotic transport	GO:0042908	0.04448435
GO:BP	pancreas development	GO:0031016	0.044846397
GO:BP	histidine metabolic process	GO:0006547	0.045882118
GO:BP	activation of immune response	GO:0002253	0.046213638
GO:BP	establishment of localization	GO:0051234	0.047806202
GO:BP	urate catabolic process	GO:0019628	0.048204113

Gene Ontology: Cellular Compartment			
GO:CC	extracellular region part	GO:0044421	1.12E-25
GO:CC	extracellular region	GO:0005576	2.79E-25
GO:CC	extracellular space	GO:0005615	6.80E-25
GO:CC	high-density lipoprotein particle	GO:0034364	8.84E-16
GO:CC	lipoprotein particle	GO:1990777	3.13E-13
GO:CC	plasma lipoprotein particle	GO:0034358	3.13E-13
GO:CC	plasma membrane region	GO:0098590	6.73E-13
GO:CC	protein-lipid complex	GO:0032994	8.56E-13
GO:CC	plasma membrane part	GO:0044459	4.44E-12
GO:CC	apical part of cell	GO:0045177	1.39E-10
GO:CC	basolateral plasma membrane	GO:0016323	1.55E-10
GO:CC	collagen-containing extracellular matrix	GO:0062023	3.86E-10
GO:CC	fibrinogen complex	GO:0005577	1.35E-09
GO:CC	endoplasmic reticulum	GO:0005783	8.82E-09
GO:CC	apical plasma membrane	GO:0016324	1.01E-08
GO:CC	extracellular matrix	GO:0031012	1.36E-08
GO:CC	intrinsic component of plasma membrane	GO:0031226	8.00E-08
GO:CC	triglyceride-rich plasma lipoprotein particle	GO:0034385	1.24E-07
GO:CC	very-low-density lipoprotein particle	GO:0034361	1.24E-07
GO:CC	integral component of plasma membrane	GO:0005887	2.90E-07
GO:CC	blood microparticle	GO:0072562	3.22E-07
GO:CC	brush border	GO:0005903	2.84682E-06
GO:CC	membrane attack complex	GO:0005579	3.62223E-06

GO:CC	platelet alpha granule	GO:0031091	1.04678E-05
GO:CC	cell surface	GO:0009986	1.80701E-05
GO:CC	endoplasmic reticulum part	GO:0044432	2.42567E-05
GO:CC	brush border membrane	GO:0031526	4.26662E-05
GO:CC	endoplasmic reticulum membrane	GO:0005789	6.17105E-05
GO:CC	nuclear outer membrane-endoplasmic reticulum membrane network	GO:0042175	6.67944E-05
GO:CC	spherical high-density lipoprotein particle	GO:0034366	8.37536E-05
GO:CC	chylomicron	GO:0042627	9.83703E-05
GO:CC	peroxisome	GO:0005777	0.000179335
GO:CC	microbody	GO:0042579	0.000179335
GO:CC	cluster of actin-based cell projections	GO:0098862	0.000234054
GO:CC	microvillus	GO:0005902	0.001777006
GO:CC	glutamatergic synapse	GO:0098978	0.006212845
GO:CC	hippocampal mossy fiber to CA3 synapse	GO:0098686	0.016330931
GO:CC	T-tubule	GO:0030315	0.018552407
GO:CC	pore complex	GO:0046930	0.019083323
GO:CC	MHC class II protein complex	GO:0042613	0.024129456
GO:CC	cell projection membrane	GO:0031253	0.029451908
GO:CC	discoidal high-density lipoprotein particle	GO:0034365	0.030899039
GO:CC	intrinsic component of postsynaptic density membrane	GO:0099146	0.034212097
GO:CC	intrinsic component of presynaptic membrane	GO:0098889	0.037117626
GO:CC	presynaptic membrane	GO:0042734	0.0395144
GO:CC	low-density lipoprotein particle	GO:0034362	0.039767519

Kyoto Encyclopedia of Genes and Genomes Pathways			
KEGG	Complement and coagulation cascades	KEGG:04610	1.51E-19
KEGG	Steroid hormone biosynthesis	KEGG:00140	1.41E-13
KEGG	Retinol metabolism	KEGG:00830	1.56E-07

KEGG	Chemical carcinogenesis	KEGG:05204	4.93E-07
KEGG	Bile secretion	KEGG:04976	4.01304E-06
KEGG	Drug metabolism - other enzymes	KEGG:00983	0.000179241
KEGG	Linoleic acid metabolism	KEGG:00591	0.000471958
KEGG	Glycine, serine and threonine metabolism	KEGG:00260	0.00064813
KEGG	Metabolism of xenobiotics by cytochrome P450	KEGG:00980	0.002500161
KEGG	Drug metabolism - cytochrome P450	KEGG:00982	0.003949089
KEGG	Arachidonic acid metabolism	KEGG:00590	0.011044643
KEGG	Platelet activation	KEGG:04611	0.01125404
KEGG	PPAR signaling pathway	KEGG:03320	0.011850295
KEGG	Primary bile acid biosynthesis	KEGG:00120	0.017092955
KEGG	Cholesterol metabolism	KEGG:04979	0.019166305
KEGG	Ascorbate and aldarate metabolism	KEGG:00053	0.023393723
KEGG	Tryptophan metabolism	KEGG:00380	0.032295018

Reactome Pathways			
REAC	Formation of Fibrin Clot (Clotting Cascade)	REAC:R-MMU-140877	4.98E-12
REAC	Platelet degranulation	REAC:R-MMU-114608	3.10E-08
REAC	Response to elevated platelet cytosolic Ca2+	REAC:R-MMU-76005	4.44E-08
REAC	Intrinsic Pathway of Fibrin Clot Formation	REAC:R-MMU-140837	1.69E-07
REAC	Metabolism of amino acids and derivatives	REAC:R-MMU-71291	1.26236E-06
REAC	Common Pathway of Fibrin Clot Formation	REAC:R-MMU-140875	1.30109E-05
REAC	Platelet activation, signaling and aggregation	REAC:R-MMU-76002	1.43987E-05
REAC	Plasma lipoprotein remodeling	REAC:R-MMU-8963899	2.21369E-05
REAC	p130Cas linkage to MAPK signaling for integrins	REAC:R-MMU-372708	3.861E-05

REAC	GRB2:SOS provides linkage to MAPK signaling for Integrins	REAC:R-MMU-354194	3.861E-05
REAC	Regulation of TLR by endogenous ligand	REAC:R-MMU-5686938	3.861E-05
REAC	Bile acid and bile salt metabolism	REAC:R-MMU-194068	8.33825E-05
REAC	Recycling of bile acids and salts	REAC:R-MMU-159418	0.000126436
REAC	Extracellular matrix organization	REAC:R-MMU-1474244	0.000135513
REAC	Hemostasis	REAC:R-MMU-109582	0.00015474
REAC	Complement cascade	REAC:R-MMU-166658	0.00023432
REAC	Integrin signaling	REAC:R-MMU-9006921	0.000274637
REAC	Integrin alphaIIb beta3 signaling	REAC:R-MMU-354192	0.000274637
REAC	Plasma lipoprotein assembly, remodeling, and clearance	REAC:R-MMU-174824	0.000292165
REAC	Platelet Aggregation (Plug Formation)	REAC:R-MMU-76009	0.000574912
REAC	MAP2K and MAPK activation	REAC:R-MMU-5674135	0.000817604
REAC	Post-translational protein phosphorylation	REAC:R-MMU-8957275	0.001093165
REAC	Synthesis of bile acids and bile salts via 7alpha-hydroxycholesterol	REAC:R-MMU-193368	0.001096415
REAC	Metabolism of fat-soluble vitamins	REAC:R-MMU-6806667	0.002070415
REAC	Regulation of Insulin-like Growth Factor (IGF) transport and uptake by Insulin-like Growth Factor Binding Proteins (IGFBPs)	REAC:R-MMU-381426	0.002126044
REAC	Regulation of Complement cascade	REAC:R-MMU-977606	0.002423426
REAC	Integrin cell surface interactions	REAC:R-MMU-216083	0.004988395

REAC	Terminal pathway of complement	REAC:R-MMU-166665	0.0082513
REAC	Retinoid metabolism and transport	REAC:R-MMU-975634	0.00924114
REAC	Plasma lipoprotein assembly	REAC:R-MMU-8963898	0.0100308
REAC	Tryptophan catabolism	REAC:R-MMU-71240	0.017197
REAC	Toll-like Receptor Cascades	REAC:R-MMU-168898	0.030901184
REAC	Metabolism of steroids	REAC:R-MMU-8957322	0.036310477
REAC	Visual phototransduction	REAC:R-MMU-2187338	0.040221958
REAC	HDL remodeling	REAC:R-MMU-8964058	0.041581529
REAC	Transport of gamma-carboxylated protein precursors from the endoplasmic reticulum to the Golgi apparatus	REAC:R-MMU-159763	0.045417385

Transcription Factor Binding Motifs			
TF	Factor: HNF-4alpha; motif: GNNCAAAGKYCANNN; match class: 0	TF:M02016_0	7.07E-07
TF	Factor: HNF-4alpha; motif: GNNCAAAGKYCANNN	TF:M02016	7.07E-07
TF	Factor: Rxrb; motif: GGGGTCAAAGGTCA; match class: 1	TF:M04493_1	9.23261E-06
TF	Factor: HNF-1beta; motif: GTTAATYATTAACY; match class: 0	TF:M03872_0	2.23106E-05
TF	Factor: HNF-1beta; motif: GTTAATYATTAACY	TF:M03872	2.23106E-05
TF	Factor: HNF4; motif: TGAMCTTTGNCCN; match class: 0	TF:M00764_0	0.000118373
TF	Factor: HNF4; motif: TGAMCTTTGNCCN	TF:M00764	0.000118373
TF	Factor: HNF-1; motif: DGTTAATKAWTNACCAM; match class: 0	TF:M00206_0	0.000122328

TF	Factor: HNF-1; motif: DGTTAATKAWTNACCAM	TF:M00206	0.000122328
TF	Factor: PPAR,; motif: RGGNCAAAGGTCA	TF:M00762	0.000626013
TF	Factor: PPAR,; motif: RGGNCAAAGGTCA; match class: 0	TF:M00762_0	0.000626013
TF	Factor: HNF-1beta; motif: GTTAATYATTAACY; match class: 1	TF:M03872_1	0.001038357
TF	Factor: PPARgamma:RXR-alpha; motif: NTRGGNCARAGGKCA; match class: 1	TF:M02262_1	0.001106799
TF	Factor: HNF-1; motif: NNNNGNTAAWNATTAACYNNN; match class: 1	TF:M01011_1	0.002165921
TF	Factor: Rxrb; motif: GGGGTCAAAGGTCA	TF:M04493	0.002244809
TF	Factor: Rxrb; motif: GGGGTCAAAGGTCA; match class: 0	TF:M04493_0	0.002244809
TF	Factor: PPAR; motif: TGACCTTTGNCCY; match class: 0	TF:M00763_0	0.002917992
TF	Factor: PPAR; motif: TGACCTTTGNCCY	TF:M00763	0.002917992
TF	Factor: CDP; motif: NNNNWGWYMAATR	TF:M04610	0.003643281
TF	Factor: CDP; motif: NNNNWGWYMAATR; match class: 0	TF:M04610_0	0.003643281
TF	Factor: HNF-1alpha; motif: GGTTAATNWTTAMCN; match class: 0	TF:M00132_0	0.004360202
TF	Factor: HNF-1alpha; motif: GGTTAATNWTTAMCN	TF:M00132	0.004360202
TF	Factor: GR; motif: NNNNNNNGKACNNNNTGTTCTNNNNNN; match class: 1	TF:M00955_1	0.004895228

TF	Factor: HNF-4A; motif: NGNNCAAAGKYCAN; match class: 1	TF:M07321_1	0.006023834
TF	Factor: HNF-1; motif: WRGTTAATNATTAACNNN	TF:M00790	0.007932802
TF	Factor: HNF-1; motif: WRGTTAATNATTAACNNN; match class: 0	TF:M00790_0	0.007932802
TF	Factor: COUP; motif: TGACCTTTGACCC; match class: 0	TF:M00765_0	0.008121556
TF	Factor: COUP; motif: TGACCTTTGACCC	TF:M00765	0.008121556
TF	Factor: EAR2; motif: YGNNCTTTGNCCTK; match class: 1	TF:M01728_1	0.008705386
TF	Factor: HNF-1; motif: WRGTTAATNATTAACNNN; match class: 1	TF:M00790_1	0.010516812
TF	Factor: HNF-4; motif: NNNRGGNCAAAGKTCANNN; match class: 1	TF:M00134_1	0.01358435
TF	Factor: EAR2; motif: YGNNCTTTGNCCTK	TF:M01728	0.014377745
TF	Factor: EAR2; motif: YGNNCTTTGNCCTK; match class: 0	TF:M01728_0	0.014377745
TF	Factor: HNF-4; motif: NNNNNNNGNNCAAAGKYCAN; match class: 1	TF:M03828_1	0.015431489
TF	Factor: HNF-4; motif: NNNRGGNCAAAGKTCANNN	TF:M00134	0.018978606
TF	Factor: HNF-4; motif: NNNRGGNCAAAGKTCANNN; match class: 0	TF:M00134_0	0.018978606
TF	Factor: RORbeta; motif: TGACCYA; match class: 1	TF:M01722_1	0.020756116
TF	Factor: HNF-4A; motif: NGNNCAAAGKYCAN	TF:M07321	0.031603367

TF	Factor: HNF-4A; motif: NGNNCAAAGKYCAN; match class: 0	TF:M07321_0	0.031603367
TF	Factor: aMEF-2; motif: CKGDYTAAAAATAACYMM	TF:M00403	0.035015042
TF	Factor: aMEF-2; motif: CKGDYTAAAAATAACYMM; match class: 0	TF:M00403_0	0.035015042
TF	Factor: HNF-4alpha; motif: GNNCAAAGKYCANNN; match class: 1	TF:M02016_1	0.044782774

Functional Enrichment in Upregulated Genes

Gene Ontology: Molecular Function			
Source	Term Name	Term ID	P(adj)
GO:MF	5S rRNA binding	GO:0008097	0.03772079

Gene Ontology: Cellular Compartment			
GO:CC	intracellular part	GO:0044424	0.009999706
GO:CC	cytosolic ribosome	GO:0022626	0.012747959
GO:CC	intracellular organelle part	GO:0044446	0.018463811
GO:CC	intracellular	GO:0005622	0.036242363
GO:CC	organelle part	GO:0044422	0.039661927
GO:CC	Cul4A-RING E3 ubiquitin ligase complex	GO:0031464	0.04328979

Transcription Factor Binding Motifs			
TF	Factor: E2F-3; motif: GGCGGGN; match class: 0	TF:M02089_0	0.000428677
TF	Factor: E2F-3; motif: GGCGGGN	TF:M02089	0.000428677
TF	Factor: AP-2gamma; motif: GCCYNNGGS; match class: 1	TF:M00470_1	0.000777899
TF	Factor: AP-2alpha; motif: GCCNNNRGS; match class: 0	TF:M00469_0	0.000910247
TF	Factor: AP-2alpha; motif: GCCNNNRGS	TF:M00469	0.000910247
TF	Factor: Mybl1; motif: NNNNNAACCGTTANNNN	TF:M02780	0.001529094
TF	Factor: Mybl1; motif: NNNNNAACCGTTANNNN; match class: 0	TF:M02780_0	0.001529094
TF	Factor: FOXN4; motif: NNWANNCGWMCGCGTCNNNNMT; match class: 1	TF:M04662_1	0.002181657
TF	Factor: E2F; motif: GGCGSG; match class: 1	TF:M00803_1	0.0034288
TF	Factor: Sp1; motif: NNGGGGCGGGGNN	TF:M00932	0.003621403
TF	Factor: Sp1; motif: NNGGGGCGGGGNN; match class: 0	TF:M00932_0	0.003621403
TF	Factor: Sp1; motif: GGGGCGGGGC	TF:M00931	0.004273116

TF	Factor: Sp1; motif: GGGGCGGGGC; match class: 0	TF:M00931_0	0.004273116
TF	Factor: E2F; motif: TTTSGCGC; match class: 0	TF:M00050_0	0.005505702
TF	Factor: E2F; motif: TTTSGCGC	TF:M00050	0.005505702
TF	Factor: SP2; motif: GNNGGGGGCGGGGSN	TF:M03807	0.007444705
TF	Factor: SP2; motif: GNNGGGGGCGGGGSN; match class: 0	TF:M03807_0	0.007444705
TF	Factor: GKLF; motif: GCCMCRCCCNNN	TF:M01588	0.00770919
TF	Factor: GKLF; motif: GCCMCRCCCNNN; match class: 0	TF:M01588_0	0.00770919
TF	Factor: AP-2alpha; motif: NGCCYSNNGSN	TF:M01857	0.009272203
TF	Factor: AP-2alpha; motif: NGCCYSNNGSN; match class: 0	TF:M01857_0	0.009272203
TF	Factor: Sp1; motif: NGGGGGCGGGGYN	TF:M00196	0.009731006
TF	Factor: Sp1; motif: NGGGGGCGGGGYN; match class: 0	TF:M00196_0	0.009731006
TF	Factor: AP-2alpha; motif: GCCNNNRGS; match class: 1	TF:M00469_1	0.009868444
TF	Factor: Sp1; motif: NGGGGCGGGGN; match class: 0	TF:M07395_0	0.01087233
TF	Factor: Sp1; motif: NGGGGCGGGGN	TF:M07395	0.01087233
TF	Factor: E2F-4; motif: GCGGGAAANA; match class: 0	TF:M02090_0	0.013169512
TF	Factor: E2F-4; motif: GCGGGAAANA	TF:M02090	0.013169512
TF	Factor: ZF5; motif: GSGCGCGR; match class: 0	TF:M00716_0	0.01460553

TF	Factor: ZF5; motif: GSGCGCGR	TF:M00716	0.01460553
TF	Factor: AP-2gamma; motif: GCCYNNGGS; match class: 0	TF:M00470_0	0.021976769
TF	Factor: AP-2gamma; motif: GCCYNNGGS	TF:M00470	0.021976769
TF	Factor: AP-2; motif: MKCCCSCNGGCG; match class: 0	TF:M00189_0	0.026224875
TF	Factor: AP-2; motif: MKCCCSCNGGCG	TF:M00189	0.026224875
TF	Factor: E2F; motif: GGCGSG	TF:M00803	0.026431078
TF	Factor: E2F; motif: GGCGSG; match class: 0	TF:M00803_0	0.026431078
TF	Factor: ZF5; motif: GSGCGCGR; match class: 1	TF:M00716_1	0.027624825
TF	Factor: Sp1; motif: CCCCGCCCCN	TF:M00933	0.044120426
TF	Factor: Sp1; motif: CCCCGCCCCN; match class: 0	TF:M00933_0	0.044120426

Appendix 4: Human Mutations Found In Predicted Driver Genes From The Input Cohort

Gene	Human Sample	Chr	Start	Tumour Allele	***in vivo* screen**
IDH1	Nakamura_BD115	2	209113113	C	**Yes**
IDH1	Nakamura_BD141	2	209113112	A	**Yes**
IDH1	Nakamura_BD167	2	209113113	A	**Yes**
IDH1	Nakamura_BD169	2	209110270	T	**Yes**
IDH1	Nakamura_BD169	2	209113113	C	**Yes**
IDH1	Nakamura_BD200	2	209113113	A	**Yes**
IDH1	Nakamura_BD212	2	209113113	A	**Yes**
IDH1	Nakamura_BD245	2	209113113	C	**Yes**
IDH1	Nakamura_BD38	2	209113113	A	**Yes**
IDH1	Nakamura_BD8	2	209113113	A	**Yes**
IDH1	Nakamura_BD81	2	209113112	A	**Yes**
IDH1	Nakamura_BD81	2	209113113	T	**Yes**
IDH1	Chan-On_p05	2	209113112	A	**Yes**
IDH1	Zou_p122	2	209113112	A	**Yes**
IDH1	Zou_p132	2	209113113	A	**Yes**
IDH1	Zou_p42	2	209113113	A	**Yes**
IDH1	Zou_p46	2	209113113	A	**Yes**
IDH1	Zou_p8_6	2	209113113	A	**Yes**
IDH1	Sia_iCC9-10	2	209113112	A	**Yes**
IDH1	TCGA_p04	2	209113113	A	**Yes**
IDH1	TCGA_p10	2	209113113	A	**Yes**
IDH1	TCGA_p23	2	209113049	A	**Yes**
IDH1	TCGA_p36	2	209113113	A	**Yes**
IDH1	TCGA_p46	2	209116300	A	**Yes**
IDH1	TCGA_p47	2	209113113	A	**Yes**
BAP1	Nakamura_BD105	3	52443569	-	**Yes**
BAP1	Nakamura_BD140	3	52441302	T	**Yes**
BAP1	Nakamura_BD141	3	52443599	-	**Yes**
BAP1	Nakamura_BD169	3	52442512	-	**Yes**
BAP1	Nakamura_BD197	3	52441990	-	**Yes**
BAP1	Nakamura_BD199	3	52441448	-	**Yes**
BAP1	Nakamura_BD21	3	52439913	A	**Yes**
BAP1	Nakamura_BD224	3	52441266	A	**Yes**
BAP1	Nakamura_BD224	3	52441337	A	**Yes**
BAP1	Nakamura_BD224	3	52441423	A	**Yes**

BAP1	Nakamura_BD226	3	52441312	-	**Yes**
BAP1	Nakamura_BD28	3	52441252	C	**Yes**
BAP1	Nakamura_BD33	3	52441333	A	**Yes**
BAP1	Nakamura_BD50	3	52437679	T	**Yes**
BAP1	Nakamura_BD79	3	52441226	A	**Yes**
BAP1	Nakamura_BD8	3	52437557	-	**Yes**
BAP1	Nakamura_BD92	3	52436633	A	**Yes**
BAP1	Chan-On_p05	3	52439880	-	**Yes**
BAP1	Chan-On_p07	3	52437650	-	**Yes**
BAP1	Chan-On_p08	3	52437802	T	**Yes**
BAP1	TCGA_p02	3	52437221	-	**Yes**
BAP1	TCGA_p03	3	52439916	G	**Yes**
BAP1	TCGA_p03	3	52439930	C	**Yes**
BAP1	TCGA_p06	3	52438377	C	**Yes**
BAP1	TCGA_p21	3	52442599	G	**Yes**
BAP1	TCGA_p23	3	52437154	A	**Yes**
BAP1	TCGA_p25	3	52440307	A	**Yes**
BAP1	TCGA_p27	3	52439230	T	**Yes**
BAP1	TCGA_p32	3	52441235	A	**Yes**
BAP1	TCGA_p47	3	52433843	T	**Yes**
ARID1A	Nakamura_BD101	1	27088022	G	**Yes**
ARID1A	Nakamura_BD104	1	27094293	T	**Yes**
ARID1A	Nakamura_BD109	1	27099285	-	**Yes**
ARID1A	Nakamura_BD120	1	27087410	-	**Yes**
ARID1A	Nakamura_BD130	1	27099947	T	**Yes**
ARID1A	Nakamura_BD141	1	27101048	T	**Yes**
ARID1A	Nakamura_BD148	1	27057916	T	**Yes**
ARID1A	Nakamura_BD154	1	27092809	T	**Yes**
ARID1A	Nakamura_BD167	1	27105978	-	**Yes**
ARID1A	Nakamura_BD169	1	27089727	T	**Yes**
ARID1A	Nakamura_BD213	1	27023717	G	**Yes**
ARID1A	Nakamura_BD213	1	27089679	-	**Yes**
ARID1A	Nakamura_BD234	1	27106648	A	**Yes**
ARID1A	Nakamura_BD237	1	27102084	A	**Yes**
ARID1A	Nakamura_BD33	1	27059264	G	**Yes**
ARID1A	Nakamura_BD58	1	27105541	T	**Yes**
ARID1A	Nakamura_BD69	1	27106106	-	**Yes**
ARID1A	Nakamura_BD81	1	27101711	T	**Yes**
ARID1A	Chan-On_p03	1	27100060	G	**Yes**
ARID1A	Chan-On_p08	1	27089777	T	**Yes**
ARID1A	Chan-On_p15	1	27092770	T	**Yes**

ARID1A	Zou_p03	1	27087458	T	**Yes**
ARID1A	Zou_p10	1	27057766	T	**Yes**
ARID1A	Zou_p108	1	27100082	A	**Yes**
ARID1A	Zou_p108	1	27100292	C	**Yes**
ARID1A	Zou_p111	1	27106504	T	**Yes**
ARID1A	Zou_p53	1	27105550	T	**Yes**
ARID1A	Zou_p56	1	27092812	T	**Yes**
ARID1A	TCGA_p01	1	27101384	-	**Yes**
ARID1A	TCGA_p03	1	27099298	-	**Yes**
ARID1A	TCGA_p14	1	27107084	T	**Yes**
ARID1A	TCGA_p16	1	27101387	-	**Yes**
ARID1A	TCGA_p24	1	27105848	-	**Yes**
ARID1A	TCGA_p27	1	27088093	C	**Yes**
ARID1A	TCGA_p41	1	27105967	T	**Yes**
ARID1A	TCGA_p41	1	27107951	-	**Yes**
ARID1A	TCGA_p42	1	27107951	-	**Yes**
ARID1A	TCGA_p45	1	27107136	A	**Yes**
ARID1A	TCGA_p46	1	27089711	-	**Yes**
KRAS	Nakamura_BD10	12	25398284	T	**Yes**
KRAS	Nakamura_BD104	12	25398284	T	**Yes**
KRAS	Nakamura_BD118	12	25380276	C	**Yes**
KRAS	Nakamura_BD118	12	25380281	G	**Yes**
KRAS	Nakamura_BD127	12	25398284	T	**Yes**
KRAS	Nakamura_BD129	12	25398281	T	**Yes**
KRAS	Nakamura_BD143	12	25380275	G	**Yes**
KRAS	Nakamura_BD148	12	25398284	A	**Yes**
KRAS	Nakamura_BD153	12	25398284	T	**Yes**
KRAS	Nakamura_BD154	12	25398284	A	**Yes**
KRAS	Nakamura_BD157	12	25398284	T	**Yes**
KRAS	Nakamura_BD18	12	25398284	T	**Yes**
KRAS	Nakamura_BD211	12	25398284	T	**Yes**
KRAS	Nakamura_BD213	12	25398284	T	**Yes**
KRAS	Nakamura_BD221	12	25398284	A	**Yes**
KRAS	Nakamura_BD229	12	25398285	A	**Yes**
KRAS	Nakamura_BD230	12	25398284	T	**Yes**
KRAS	Nakamura_BD233	12	25378562	T	**Yes**
KRAS	Nakamura_BD234	12	25398266	C	**Yes**
KRAS	Nakamura_BD243	12	25398284	T	**Yes**
KRAS	Nakamura_BD247	12	25398284	T	**Yes**
KRAS	Nakamura_BD25	12	25398284	A	**Yes**
KRAS	Nakamura_BD29	12	25398284	T	**Yes**

KRAS	Nakamura_BD30	12	25398284	A	**Yes**
KRAS	Nakamura_BD47	12	25380275	G	**Yes**
KRAS	Nakamura_BD5	12	25398284	T	**Yes**
KRAS	Nakamura_BD66	12	25398284	T	**Yes**
KRAS	Nakamura_BD74	12	25398284	G	**Yes**
KRAS	Nakamura_BD86	12	25398284	A	**Yes**
KRAS	Nakamura_BD88	12	25398284	T	**Yes**
KRAS	Zou_p03	12	25398284	T	**Yes**
KRAS	Zou_p07	12	25398284	T	**Yes**
KRAS	Zou_p107	12	25398284	A	**Yes**
KRAS	Zou_p108	12	25398284	A	**Yes**
KRAS	Zou_p111	12	25398284	T	**Yes**
KRAS	Zou_p120	12	25398285	A	**Yes**
KRAS	Zou_p129	12	25398284	T	**Yes**
KRAS	Zou_p133	12	25398285	A	**Yes**
KRAS	Zou_p134	12	25398284	G	**Yes**
KRAS	Zou_p137	12	25398284	A	**Yes**
KRAS	Zou_p144	12	25398284	A	**Yes**
KRAS	Zou_p17	12	25398284	T	**Yes**
KRAS	Zou_p29	12	25398284	A	**Yes**
KRAS	Zou_p33	12	25398284	A	**Yes**
KRAS	Zou_p46	12	25398284	T	**Yes**
KRAS	Zou_p59	12	25398284	T	**Yes**
KRAS	Zou_p66	12	25398284	G	**Yes**
KRAS	TCGA_p25	12	25359465	-	**Yes**
KRAS	TCGA_p26	12	25380277	T	**Yes**
KRAS	TCGA_p26	12	25380278	T	**Yes**
KRAS	TCGA_p32	12	25359046	A	**Yes**
KRAS	TCGA_p41	12	25398284	T	**Yes**
TP53	Nakamura_BD111	17	7577085	T	**Yes**
TP53	Nakamura_BD120	17	7577538	T	**Yes**
TP53	Nakamura_BD129	17	7577580	C	**Yes**
TP53	Nakamura_BD130	17	7578212	A	**Yes**
TP53	Nakamura_BD134	17	7577538	A	**Yes**
TP53	Nakamura_BD140	17	7578517	A	**Yes**
TP53	Nakamura_BD143	17	7578550	A	**Yes**
TP53	Nakamura_BD154	17	7576852	A	**Yes**
TP53	Nakamura_BD18	17	7577022	A	**Yes**
TP53	Nakamura_BD203	17	7577534	A	**Yes**
TP53	Nakamura_BD211	17	7578191	T	**Yes**
TP53	Nakamura_BD213	17	7577106	T	**Yes**

TP53	Nakamura_BD214	17	7577557	C	**Yes**
TP53	Nakamura_BD220	17	7578550	A	**Yes**
TP53	Nakamura_BD227	17	7578265	T	**Yes**
TP53	Nakamura_BD229	17	7577539	C	**Yes**
TP53	Nakamura_BD23	17	7578190	C	**Yes**
TP53	Nakamura_BD231	17	7578271	C	**Yes**
TP53	Nakamura_BD234	17	7579472	G	**Yes**
TP53	Nakamura_BD243	17	7578234	-	**Yes**
TP53	Nakamura_BD244	17	7577094	A	**Yes**
TP53	Nakamura_BD25	17	7578406	T	**Yes**
TP53	Nakamura_BD29	17	7578265	G	**Yes**
TP53	Nakamura_BD38	17	7578369	G	**Yes**
TP53	Nakamura_BD74	17	7578529	C	**Yes**
TP53	Nakamura_BD82	17	7578177	T	**Yes**
TP53	Nakamura_BD86	17	7576855	A	**Yes**
TP53	Nakamura_BD88	17	7576926	A	**Yes**
TP53	Nakamura_BD88	17	7576927	A	**Yes**
TP53	Chan-On_p02	17	7578457	T	**Yes**
TP53	Zou_p05	17	7577568	A	**Yes**
TP53	Zou_p08	17	7578493	T	**Yes**
TP53	Zou_p113	17	7579406	T	**Yes**
TP53	Zou_p121	17	7578440	A	**Yes**
TP53	Zou_p127	17	7577498	A	**Yes**
TP53	Zou_p128	17	7573996	C	**Yes**
TP53	Zou_p131	17	7578239	A	**Yes**
TP53	Zou_p134	17	7574017	A	**Yes**
TP53	Zou_p135	17	7577018	A	**Yes**
TP53	Zou_p140	17	7579311	A	**Yes**
TP53	Zou_p143	17	7578530	G	**Yes**
TP53	Zou_p145	17	7578235	C	**Yes**
TP53	Zou_p16	17	7574035	A	**Yes**
TP53	Zou_p17	17	7578404	T	**Yes**
TP53	Zou_p34	17	7577509	G	**Yes**
TP53	Zou_p44	17	7578177	T	**Yes**
TP53	Zou_p58	17	7578224	A	**Yes**
TP53	Zou_p66	17	7578535	C	**Yes**
TP53	Zou_p8_4	17	7578410	A	**Yes**
TP53	Zou_p87	17	7578226	A	**Yes**
TP53	Zou_p95	17	7576927	A	**Yes**
TP53	Zou_p95	17	7578419	A	**Yes**
TP53	TCGA_p05	17	7578509	G	**Yes**

TP53	TCGA_p12	17	7592560	T	**Yes**
TP53	TCGA_p22	17	7574005	T	**Yes**
PBRM1	Nakamura_BD12	3	52595944	-	**Yes**
PBRM1	Nakamura_BD147	3	52696293	T	**Yes**
PBRM1	Nakamura_BD167	3	52668672	T	**Yes**
PBRM1	Nakamura_BD211	3	52598156	C	**Yes**
PBRM1	Nakamura_BD245	3	52643728	A	**Yes**
PBRM1	Nakamura_BD247	3	52685758	T	**Yes**
PBRM1	Nakamura_BD31	3	52675988	A	**Yes**
PBRM1	Nakamura_BD33	3	52661291	T	**Yes**
PBRM1	Zou_p8_1	3	52661288	G	**Yes**
PBRM1	TCGA_p02	3	52643837	T	**Yes**
PBRM1	TCGA_p02	3	52662939	A	**Yes**
PBRM1	TCGA_p03	3	52592205	A	**Yes**
PBRM1	TCGA_p03	3	52651496	A	**Yes**
PBRM1	TCGA_p05	3	52621439	AAT	**Yes**
PBRM1	TCGA_p06	3	52661362	A	**Yes**
PBRM1	TCGA_p07	3	52713533	T	**Yes**
PBRM1	TCGA_p09	3	52682330	-	**Yes**
PBRM1	TCGA_p09	3	52682473	T	**Yes**
PBRM1	TCGA_p14	3	52651334	A	**Yes**
PBRM1	TCGA_p17	3	52649497	G	**Yes**
PBRM1	TCGA_p20	3	52621364	T	**Yes**
PBRM1	TCGA_p33	3	52620649	-	**Yes**
PBRM1	TCGA_p37	3	52592236	-	**Yes**
PBRM1	TCGA_p38	3	52678790	A	**Yes**
PBRM1	TCGA_p45	3	52668817	A	**Yes**
IDH2	Nakamura_BD117	15	90631838	T	**Yes**
IDH2	Nakamura_BD149	15	90631838	T	**Yes**
IDH2	Nakamura_BD219	15	90631838	T	**Yes**
IDH2	Chan-On_p07	15	90622990	T	**Yes**
IDH2	TCGA_p09	15	90631838	T	**Yes**
IDH2	TCGA_p23	15	90631837	A	**Yes**
IDH2	TCGA_p33	15	90631838	T	**Yes**
IDH2	TCGA_p36	15	90634599	T	**Yes**

References

1. Marin, J. J. G. *et al.* Chemoresistance and chemosensitization in cholangiocarcinoma. *Biochim. Biophys. acta. Mol. basis Dis.* **1864**, 1444–1453 (2018).

2. Wang, K., Zhang, H., Xia, Y., Liu, J. & Shen, F. Surgical options for intrahepatic cholangiocarcinoma. *Hepatobiliary Surg. Nutr.* **6**, 79–90 (2017).

3. Hu, L.-S. *et al.* Impact of microvascular invasion on clinical outcomes after curative-intent resection for intrahepatic cholangiocarcinoma. *J. Surg. Oncol.* **119**, 21–29 (2019).

4. Jutric, Z. *et al.* Impact of lymph node status in patients with intrahepatic cholangiocarcinoma treated by major hepatectomy: a review of the National Cancer Database. *HPB (Oxford).* **18**, 79–87 (2016).

5. Yamamoto, M., Takasaki, K. & Yoshikawa, T. Lymph node metastasis in intrahepatic cholangiocarcinoma. *Jpn. J. Clin. Oncol.* **29**, 147–150 (1999).

6. Yamamoto, M. *et al.* A Long-term Survivor of Intrahepatic Cholangiocarcinoma with Lymph Node Metastasis: a Case Report. *Jpn. J. Clin. Oncol.* **32**, 206–209 (2002).

7. Hoyos, S., Navas, M. C., Restrepo, J. C. & Botero, R. C. Current controversies in cholangiocarcinoma. *Biochim. Biophys. Acta - Mol. Basis Dis.* (2018) doi:10.1016/j.bbadis.2017.07.027.

8. Huang, Y., Li, X. & Zhao, Y. Progression of targeted therapy in advanced cholangiocarcinoma. *Chin. J. Cancer Res.* **27**, 122–127 (2015).

9. Simile, M. M. *et al.* Targeted Therapies in Cholangiocarcinoma: Emerging Evidence from Clinical Trials. *Medicina (Kaunas).* **55**, 42 (2019).

10. Verdaguer, H. *et al.* Impact of cholangiocarcinoma (CC) molecular heterogeneity on outcome during first-line chemotherapy and access to targeted therapies in early clinical trials (CT). *J. Clin. Oncol.* **36**, 4091 (2018).

11. Brandi, G., Farioli, A., Astolfi, A., Biasco, G. & Tavolari, S. Genetic heterogeneity in cholangiocarcinoma: a major challenge for targeted therapies. *Oncotarget* **6**, 14744–53 (2015).

12. Walter, D. *et al.* Intratumoral heterogeneity of intrahepatic cholangiocarcinoma. *Oncotarget* **8**, 14957–14968 (2017).

13. Massarweh, N. N. & El-Serag, H. B. Epidemiology of Hepatocellular Carcinoma and Intrahepatic Cholangiocarcinoma. *Cancer Control* **24**, 1–11 (2017).

14. Groen, P. C. De, Gores, G. J., Larusso, N. F., Gunderson, L. L. & Nagorney, D. M. Biliary Tract Cancers. *N. Engl. J. Med.* (1999).

15. Khan, S. A., Thomas, H. C., Davidson, B. R. & Taylor-robinson, S. D. Cholangiocarcinoma. **366**, (2005).

16. Taylor-Robinson, S. D. *et al.* Increase in mortality rates from intrahepatic

cholangiocarcinoma in England and Wales 1968-1998. *Gut* **48**, 816–820 (2001).

17. Kendall, T. *et al.* Anatomical, histomorphological and molecular classification of cholangiocarcinoma. *Liver Int.* **39**, 7–18 (2019).

18. Lim, J. H. Cholangiocarcinoma: Morphologic Classification According to Growth Pattern and Imaging Findings. *Am. J. Roentgenol.* **181**, 819–827 (2003).

19. Aishima, S. & Oda, Y. Pathogenesis and classification of intrahepatic cholangiocarcinoma: Different characters of perihilar large duct type versus peripheral small duct type. *J. Hepatobiliary. Pancreat. Sci.* **22**, 94–100 (2015).

20. Nakanuma, Y. *et al.* Pathological classification of intrahepatic cholangiocarcinoma based on a new concept. *World J. Hepatol.* **2**, 419–427 (2010).

21. Banales, J. M. *et al.* Expert consensus document: Cholangiocarcinoma: current knowledge and future perspectives consensus statement from the European Network for the Study of Cholangiocarcinoma (ENS-CCA). *Nat. Rev. Gastroenterol. Hepatol.* **13**, 261–80 (2016).

22. Kirstein, M. M. & Vogel, A. Epidemiology and Risk Factors of Cholangiocarcinoma. *Visc. Med.* **32**, 395–400 (2016).

23. Khan, S. A., Toledano, M. B. & Taylor-Robinson, S. D. Epidemiology, risk factors, and pathogenesis of cholangiocarcinoma. *Hpb* **10**, 77–82 (2008).

24. Khan, S. A. *et al.* Rising trends in cholangiocarcinoma: Is the ICD classification system misleading us? *J. Hepatol.* **56**, 848–854 (2012).

25. Al-Bahrani, R., Abuetabh, Y., Zeitouni, N. & Sergi, C. Cholangiocarcinoma: risk factors, environmental influences and oncogenesis. *Ann. Clin. Lab. Sci.* **43**, 195–210 (2013).

26. Khan, S. A., Tavolari, S. & Brandi, G. Cholangiocarcinoma: Epidemiology and risk factors. *Liver Int.* **39**, 19–31 (2019).

27. Fried, B., Graczyk, T. K. & Tamang, L. Food-borne intestinal trematodiases in humans. *Parasitol. Res.* **93**, 159–170 (2004).

28. Kaewpitoon, N., Kaewpitoon, S. J., Pengsaa, P. & Sripa, B. Opisthorchis viverrini: The carcinogenic human liver fluke. *World J. Gastroenterol.* **14**, 666–674 (2008).

29. Khuntikeo, N. *et al. The Socioeconomic Burden of Cholangiocarcinoma Associated With Opisthorchis viverrini Sensu Lato Infection in Northeast Thailand: A Preliminary Analysis*. *Advances in Parasitology* vol. 102 (Elsevier Ltd, 2018).

30. Sithithaworn, P. *et al.* The current status of opisthorchiasis and clonorchiasis in the Mekong Basin. *Parasitol. Int.* **61**, 10–16 (2012).

31. Thammatacharee, N., Suphanchaimat, R., Wisaijohn, T., Limwattananon, S. & Putthasri, W. Attitudes toward working in rural areas of Thai medical, dental and pharmacy new graduates in 2012: A cross-sectional survey. *Hum. Resour. Health* **11**, 1–10 (2013).

32. Songserm, N., Prasongwattana, J., Sithithaworn, P., Sripa, B. & Pipitgool, V.

Songserm, Pipitgool - CCA in experimental hamsters with long-standing OV infection 2009. **10**, 299–302 (2009).

33. Pairojkul, C. *et al.* Multistage carcinogenesis of liver-fluke-associated cholangiocarcinoma in Thailand. *Princess Takamatsu Symp.* **22**, 77—86 (1991).

34. Isvilanonda, S., Ahmad, A. & Hossain, M. Recent Changes in Thailand's Rural Economy: Evidence from Six Villages. *Econ. Polit. Wkly.* **35**, 4644–4649 (2000).

35. Krongkaew, M. Thailand's Internationalization and Its Rural Sector. *ASEAN Econ. Bull.* **11**, 306–319 (1995).

36. Barnaud, C., Trébuil, G., Dufumier, M. & Suphanchaimat, N. Rural Poverty and Diversification of Farming Systems in Upper Northeast Thailand. *Moussons* **9–10**, 157–187 (2006).

37. Suk, W. A., Bhudhisawasdi, V. & Ruchirawat, M. The Curious Case of Cholangiocarcinoma: Opportunities for Environmental Health Scientists to Learn about a Complex Disease. *J. Environ. Public Health* **2018**, (2018).

38. Steele, J. A. *et al.* Thinking beyond Opisthorchis viverrini for risk of cholangiocarcinoma in the lower Mekong region: A systematic review and meta-analysis. *Infect. Dis. Poverty* **7**, 1–13 (2018).

39. Bernstein, H., Bernstein, C., Payne, C. M., Dvorakova, K. & Garewal, H. Bile acids as carcinogens in human gastrointestinal cancers. *Mutat. Res.* **589**, 47–65 (2005).

40. Debruyne, P. R. *et al.* The role of bile acids in carcinogenesis. *Mutat. Res. Mol. Mech. Mutagen.* **480–481**, 359–369 (2001).

41. Park, J. Y. *et al.* Bile acid analysis in biliary tract cancer. *Yonsei Med. J.* **47**, 817–825 (2006).

42. Jusakul, A. *et al.* Identification of biliary bile acids in patients with benign biliary diseases, hepatocellular carcinoma and cholangiocarcinoma. *Asian Pac. J. Cancer Prev.* **13 Suppl**, 77–82 (2012).

43. Lozano, E. *et al.* Cocarcinogenic effects of intrahepatic bile acid accumulation in cholangiocarcinoma development. *Mol. Cancer Res.* **12**, 91–100 (2014).

44. Melum, E. *et al.* Cholangiocarcinoma in primary sclerosing cholangitis is associated with NKG2D polymorphisms. *Hepatology* **47**, 90–96 (2008).

45. Wadsworth, C. A. *et al.* Polymorphisms in Natural Killer Cell Receptor Protein 2D (NKG2D) as a Risk Factor for Cholangiocarcinoma. *J. Clin. Exp. Hepatol.* **9**, 171–175 (2019).

46. Marquardt, J. U., Andersen, J. B. & Thorgeirsson, S. S. Functional and genetic deconstruction of the cellular origin in liver cancer. *Nat. Rev. Cancer* **15**, 653–67 (2015).

47. Chaisaingmongkol, J. *et al.* Common Molecular Subtypes Among Asian Hepatocellular Carcinoma and Cholangiocarcinoma. *Cancer Cell* **32**, 57-70.e3 (2017).

48. Seok, J. Y. *et al.* A fibrous stromal component in hepatocellular carcinoma reveals a cholangiocarcinoma-like gene expression trait and epithelial-mesenchymal transition. *Hepatology* **55**, 1776–1786 (2012).

49. Woo, H. G. *et al.* Identification of a cholangiocarcinoma-like gene expression trait in hepatocellular carcinoma. *Cancer Res.* **70**, 3034–3041 (2010).

50. Guest, R. V. *et al.* Cell lineage tracing reveals a biliary origin of intrahepatic cholangiocarcinoma. *Cancer Res.* **74**, 1005–1010 (2014).

51. Tanimizu, N. *et al.* Progressive induction of hepatocyte progenitor cells in chronically injured liver. *Sci. Rep.* **7**, 39990 (2017).

52. Sritananuwat, P., Sueangoen, N., Thummarati, P., Islam, K. & Suthiphongchai, T. Blocking ERK1/2 signaling impairs TGF-β1 tumor promoting function but enhances its tumor suppressing role in intrahepatic cholangiocarcinoma cells. *Cancer Cell Int.* **17**, 85 (2017).

53. Huang, C.-K. *et al.* Expression of transforming growth factor β1 promotes cholangiocarcinoma development and progression. *Cancer Lett.* **380**, 153–162 (2016).

54. Lustri, A. M. *et al.* TGF-β signaling is an effective target to impair survival and induce apoptosis of human cholangiocarcinoma cells: A study on human primary cell cultures. *PLoS One* **12**, e0183932 (2017).

55. Mu, X. *et al.* Epithelial Transforming Growth Factor-β Signaling Does Not Contribute to Liver Fibrosis but Protects Mice From Cholangiocarcinoma. *Gastroenterology* **150**, 720–733 (2016).

56. Yothaisong, S. *et al.* Increased activation of PI3K/AKT signaling pathway is associated with cholangiocarcinoma metastasis and PI3K/mTOR inhibition presents a possible therapeutic strategy. *Tumour Biol.* **34**, 3637–3648 (2013).

57. Wu, C.-E., Chen, M.-H. & Yeh, C.-N. mTOR Inhibitors in Advanced Biliary Tract Cancers. *Int. J. Mol. Sci.* **20**, 500 (2019).

58. Corti, F. *et al.* Targeting the PI3K/AKT/mTOR pathway in biliary tract cancers: A review of current evidences and future perspectives. *Cancer Treat. Rev.* **72**, 45–55 (2019).

59. Leelawat, K., Narong, S., Udomchaiprasertkul, W., Leelawat, S. & Tungpradubkul, S. Inhibition of PI3K increases oxaliplatin sensitivity in cholangiocarcinoma cells. *Cancer Cell Int.* **9**, 3 (2009).

60. Kim, Y. *et al.* Hedgehog signaling between cancer cells and hepatic stellate cells in promoting cholangiocarcinoma. *Ann. Surg. Oncol.* **21**, 2684–2698 (2014).

61. Jinawath, A., Akiyama, Y., Sripa, B. & Yuasa, Y. Dual blockade of the Hedgehog and ERK1/2 pathways coordinately decreases proliferation and survival of cholangiocarcinoma cells. *J. Cancer Res. Clin. Oncol.* **133**, 271–278 (2007).

62. Gao, L., Zhang, Z., Zhang, P., Yu, M. & Yang, T. Role of canonical Hedgehog signaling pathway in liver. *Int. J. Biol. Sci.* **14**, 1636–1644 (2018).

63. Senggunprai, L., Kukongviriyapan, V., Prawan, A. & Kukongviriyapan, U. Quercetin and EGCG exhibit chemopreventive effects in cholangiocarcinoma cells via suppression of JAK/STAT signaling pathway. *Phytother. Res.* **28**, 841–848 (2014).

64. Cao, Y., Xia, G., Guo, J. & Cheng, K. MiR-221 affects cholangiocarcinoma cell proliferation and apoptosis by targeting the SOCS3 / JAK-STAT signaling pathway. **12**, 12228–12235 (2019).

65. Yang, J. *et al.* Cholangiocarcinoma cells secrete soluble factors that activate Jak/STAT signal transduction and promote MDSC expansion. *J. Immunother. Cancer* **2**, P271 (2014).

66. Smirnova, O. V, Ostroukhova, T. Y. & Bogorad, R. L. JAK-STAT pathway in carcinogenesis: is it relevant to cholangiocarcinoma progression? *World J. Gastroenterol.* **13**, 6478–6491 (2007).

67. Hilman, D. & Gat, U. The evolutionary history of YAP and the Hippo/YAP pathway. *Mol. Biol. Evol.* **28**, 2403–2417 (2011).

68. Bossuyt, W. *et al.* An evolutionary shift in the regulation of the Hippo pathway between mice and flies. *Oncogene* **33**, 1218–1228 (2014).

69. Lin, C., Yao, E. & Chuang, P.-T. A conserved MST1/2–YAP axis mediates Hippo signaling during lung growth. *Dev. Biol.* **403**, 101–113 (2015).

70. Qin, F., Tian, J., Zhou, D. & Chen, L. Mst1 and Mst2 kinases: regulations and diseases. *Cell Biosci.* **3**, 31 (2013).

71. Chan, E. H. Y. *et al.* The Ste20-like kinase Mst2 activates the human large tumor suppressor kinase Lats1. *Oncogene* **24**, 2076–2086 (2005).

72. Kim, W. & Jho, E. H. The history and regulatory mechanism of the Hippo pathway. *BMB Rep.* **51**, 106–118 (2018).

73. Kim, M.-K., Jang, J.-W. & Bae, S.-C. DNA binding partners of YAP/TAZ. *BMB Rep.* **51**, 126–133 (2018).

74. Zhao, B. *et al.* TEAD mediates YAP-dependent gene induction and growth control. *Genes Dev.* **22**, 1962–1971 (2008).

75. Huang, J., Wu, S., Barrera, J., Matthews, K. & Pan, D. The Hippo signaling pathway coordinately regulates cell proliferation and apoptosis by inactivating Yorkie, the Drosophila Homolog of YAP. *Cell* **122**, 421–434 (2005).

76. Yu, F.-X. *et al.* Regulation of the Hippo-YAP Pathway by G-Protein-Coupled Receptor Signaling. *Cell* **150**, 780–791 (2012).

77. Boggiano, J. C., Vanderzalm, P. J. & Fehon, R. G. Tao-1 phosphorylates Hippo/MST kinases to regulate the Hippo-Salvador-Warts tumor suppressor pathway. *Dev. Cell* **21**, 888–895 (2011).

78. Martin, D. *et al.* Assembly and activation of the Hippo signalome by FAT1 tumor suppressor. *Nat. Commun.* **9**, 2372 (2018).

79. Meng, Z. *et al.* MAP4K family kinases act in parallel to MST1/2 to activate LATS1/2 in the Hippo pathway. *Nat. Commun.* **6**, 8357 (2015).

80. Badouel, C. *et al.* The FERM-domain protein Expanded regulates Hippo pathway activity via direct interactions with the transcriptional activator Yorkie. *Dev. Cell* **16**, 411–420 (2009).

81. Li, Y., Wei, Z., Zhang, J., Yang, Z. & Zhang, M. Structural basis of the binding of Merlin FERM domain to the E3 ubiquitin ligase substrate adaptor DCAF1. *J. Biol. Chem.* **289**, 14674–14681 (2014).

82. Kang, B. S., Cooper, D. R., Devedjiev, Y., Derewenda, U. & Derewenda, Z. S. The structure of the FERM domain of merlin, the neurofibromatosis type 2 gene product. *Acta Crystallogr. D. Biol. Crystallogr.* **58**, 381–391 (2002).

83. Evans, Dg. R. Neurofibromatosis type 2 (NF2): A clinical and molecular review. *Orphanet J. Rare Dis.* **4**, 16 (2009).

84. Grijalva, J. L. *et al.* Dynamic alterations in Hippo signaling pathway and YAP activation during liver regeneration. *Am. J. Physiol. Gastrointest. Liver Physiol.* **307**, G196–G204 (2014).

85. Wu, H. *et al.* The Ets transcription factor GABP is a component of the hippo pathway essential for growth and antioxidant defense. *Cell Rep.* **3**, 1663–1677 (2013).

86. Manmadhan, S. & Ehmer, U. Hippo Signaling in the Liver - A Long and Ever-Expanding Story. *Front. cell Dev. Biol.* **7**, 33 (2019).

87. Sadler, K. C., Amsterdam, A., Soroka, C., Boyer, J. & Hopkins, N. A genetic screen in zebrafish identifies the mutants vps18, nf2 and foie gras as models of liver disease. *Development* **132**, 3561–3572 (2005).

88. Benhamouche, S. *et al.* Nf2/Merlin controls progenitor homeostasis and tumorigenesis in the liver. *Genes Dev.* **24**, 1718–1730 (2010).

89. Yimlamai, D. *et al.* Hippo pathway activity influences liver cell fate. *Cell* **157**, 1324–1338 (2014).

90. Zhang, N. *et al.* Molecular alterations of the NF2 gene in hepatocellular carcinoma and intrahepatic cholangiocarcinoma. *Oncol. Rep.* **38**, 3650–3658 (2017).

91. Andersson, E. R., Sandberg, R. & Lendahl, U. Notch signaling: Simplicity in design, versatility in function. *Development* **138**, 3593–3612 (2011).

92. Raynaud, P., Carpentier, R., Antoniou, A. & Lemaigre, F. P. Biliary differentiation and bile duct morphogenesis in development and disease. *Int. J. Biochem. Cell Biol.* **43**, 245–256 (2011).

93. Antoniou, A. *et al.* Intrahepatic Bile Ducts Develop According to a New Mode of Tubulogenesis Regulated by the Transcription Factor SOX9. *Gastroenterology* **136**, 2325–2333 (2009).

94. Schaub, J. R. *et al.* De novo formation of the biliary system by TGFβ-mediated

hepatocyte transdifferentiation. *Nature* **557**, 247–251 (2018).

95. McDaniell, R. *et al.* NOTCH2 mutations cause Alagille syndrome, a heterogeneous disorder of the notch signaling pathway. *Am. J. Hum. Genet.* **79**, 169–173 (2006).

96. Köhler, C. *et al.* Expression of Notch-1 and its ligand Jagged-1 in rat liver during liver regeneration. *Hepatology* **39**, 1056–1065 (2004).

97. Croquelois, A. *et al.* Inducible inactivation of Notch1 causes nodular regenerative hyperplasia in mice. *Hepatology* **41**, 487–496 (2005).

98. Dill, M. T. *et al.* Disruption of Notch1 induces vascular remodeling, intussusceptive angiogenesis, and angiosarcomas in livers of mice. *Gastroenterology* **142**, 967-977.e2 (2012).

99. Fan, B. *et al.* Cholangiocarcinomas can originate from hepatocytes in mice. *J. Clin. Invest.* **122**, 2911–2915 (2012).

100. Sekiya, S. & Suzuki, A. Intrahepatic cholangiocarcinoma can arise from Notch-mediated conversion of hepatocytes. *J. Clin. Invest.* **122**, 3914–3918 (2012).

101. Gao, J. *et al.* RUNX3 directly interacts with intracellular domain of Notch1 and suppresses Notch signaling in hepatocellular carcinoma cells. *Exp. Cell Res.* **316**, 149–157 (2010).

102. Hayashi, Y., Osanai, M. & Lee, G.-H. NOTCH2 signaling confers immature morphology and aggressiveness in human hepatocellular carcinoma cells. *Oncol. Rep.* **34**, 1650–1658 (2015).

103. O'Rourke, C. J. *et al.* Identification of a Pan-Gamma-Secretase Inhibitor Response Signature for Notch-Driven Cholangiocarcinoma. *Hepatology* **0**, 1–18 (2019).

104. Guest, R. V. *et al.* Notch3 drives development and progression of cholangiocarcinoma. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 12250–12255 (2016).

105. Janku, F., Yap, T. A. & Meric-Bernstam, F. Targeting the PI3K pathway in cancer: are we making headway? *Nat. Rev. Clin. Oncol.* **15**, 273–291 (2018).

106. Janku, F. *et al.* PIK3CA mutation H1047R is associated with response to PI3K/AKT/mTOR signaling pathway inhibitors in early-phase clinical trials. *Cancer Res.* **73**, 276–284 (2013).

107. Ackers, I. & Malgor, R. Interrelationship of canonical and non-canonical Wnt signalling pathways in chronic metabolic diseases. *Diabetes Vasc. Dis. Res.* **15**, 3–13 (2018).

108. Grumolato, L. *et al.* Canonical and noncanonical Wnts use a common mechanism to activate completely unrelated coreceptors. *Genes Dev.* **24**, 2517–2530 (2010).

109. Chae, W.-J. & Bothwell, A. L. M. Canonical and Non-Canonical Wnt Signaling in Immune Cells. *Trends Immunol.* **39**, 830–847 (2018).

110. Liu, J. *et al.* Siah-1 Mediates a Novel β-Catenin Degradation Pathway Linking p53 to the Adenomatous Polyposis Coli Protein. *Mol. Cell* **7**, 927–936 (2001).

111. Matsuzawa, S. & Reed, J. C. Siah-1, SIP, and Ebi Collaborate in a Novel Pathway

for β-Catenin Degradation Linked to p53 Responses. *Mol. Cell* 7, 915–926 (2001).

112. Valenta, T., Hausmann, G. & Basler, K. The many faces and functions of β-catenin. *EMBO J.* **31**, 2714–2736 (2012).

113. Cha, B. *et al.* Mechanotransduction activates canonical Wnt/β-catenin signaling to promote lymphatic vascular patterning and the development of lymphatic and lymphovenous valves. *Genes Dev.* **30**, 1454–1469 (2016).

114. Pukhlyakova, E., Aman, A. J., Elsayad, K. & Technau, U. β-Catenin–dependent mechanotransduction dates back to the common ancestor of Cnidaria and Bilateria. *Proc. Natl. Acad. Sci.* **115**, 6231 LP – 6236 (2018).

115. Brunet, T. *et al.* Evolutionary conservation of early mesoderm specification by mechanotransduction in Bilateria. *Nat. Commun.* **4**, 2821 (2013).

116. Bauer, A. *et al.* Pontin52 and reptin52 function as antagonistic regulators of beta-catenin signalling activity. *EMBO J.* **19**, 6121–6130 (2000).

117. Novak, A. & Dedhar, S. Signaling through beta-catenin and Lef/Tcf. *Cell. Mol. Life Sci.* **56**, 523–537 (1999).

118. Xu, W. & Kimelman, D. Mechanistic insights from structural studies of β-catenin and its binding partners. *J. Cell Sci.* **120**, 3337 LP – 3344 (2007).

119. Kimelman, D. & Xu, W. β-Catenin destruction complex: insights and questions from a structural perspective. *Oncogene* **25**, 7482–7491 (2006).

120. Stamos, J. L. & Weis, W. I. The β-catenin destruction complex. *Cold Spring Harb. Perspect. Biol.* **5**, a007898–a007898 (2013).

121. Winston, J. T. *et al.* The SCFbeta-TRCP-ubiquitin ligase complex associates specifically with phosphorylated destruction motifs in IkappaBalpha and beta-catenin and stimulates IkappaBalpha ubiquitination in vitro. *Genes Dev.* **13**, 270–283 (1999).

122. Mikels, A. J. & Nusse, R. Wnts as ligands: processing, secretion and reception. *Oncogene* **25**, 7461–7468 (2006).

123. Ciruna, B., Jenny, A., Lee, D., Mlodzik, M. & Schier, A. F. Planar cell polarity signalling couples cell division and morphogenesis during neurulation. *Nature* **439**, 220–224 (2006).

124. Wang, J. *et al.* Dishevelled genes mediate a conserved mammalian PCP pathway to regulate convergent extension during neurulation. *Development* **133**, 1767 LP – 1778 (2006).

125. Yang, Y. & Mlodzik, M. Wnt-Frizzled/Planar Cell Polarity Signaling: Cellular Orientation by Facing the Wind (Wnt). *Annu. Rev. Cell Dev. Biol.* **31**, 623–646 (2015).

126. Minegishi, K. *et al.* A Wnt5 Activity Asymmetry and Intercellular Signaling via PCP Proteins Polarize Node Cells for Left-Right Symmetry Breaking. *Dev. Cell* **40**, 439-452.e4 (2017).

127. De, A. Wnt/Ca2+ signaling pathway: a brief overview. *Acta Biochim. Biophys. Sin. (Shanghai).* **43**, 745–756 (2011).

128. Perugorria, M. J. *et al.* Wnt–β-catenin signalling in liver development, health and disease. *Nat. Rev. Gastroenterol. Hepatol.* **16**, 121–136 (2019).

129. Miller, M. F. *et al.* Wnt ligands signal in a cooperative manner to promote foregut organogenesis. *Proc. Natl. Acad. Sci.* **109**, 15348 LP – 15353 (2012).

130. Zorn. Liver Development. *StemBook* 1–26 (2008) doi:10.3824/stembook.1.25.1.

131. Decaens, T. *et al.* Stabilization of β-catenin affects mouse embryonic liver growth and hepatoblast fate. *Hepatology* **47**, 247–258 (2008).

132. Cui, S., Capecci, L. M. & Matthews, R. P. Disruption of planar cell polarity activity leads to developmental biliary defects. *Dev. Biol.* **351**, 229–241 (2011).

133. Pez, F. *et al.* Wnt signaling and hepatocarcinogenesis: Molecular targets for the development of innovative anticancer drugs. *J. Hepatol.* **59**, 1107–1117 (2013).

134. Zheng, Z. *et al.* MicroRNA-452 promotes stem-like cells of hepatocellular carcinoma by inhibiting Sox7 involving Wnt/β-catenin signaling pathway. *Oncotarget* **7**, 28000–28012 (2016).

135. Tang, J. *et al.* miR-612 suppresses the stemness of liver cancer via Wnt/β-catenin signaling. *Biochem. Biophys. Res. Commun.* **447**, 210–215 (2014).

136. Tripathy, A. *et al.* The molecular connection of histopathological heterogeneity in hepatocellular carcinoma: A role of Wnt and Hedgehog signaling pathways. *PLoS One* **13**, e0208194 (2018).

137. Boulter, L. *et al.* WNT signaling drives cholangiocarcinoma growth and can be pharmacologically inhibited. *J. Clin. Invest.* **125**, 1269–1285 (2015).

138. Wang, Y. *et al.* SRC-like adaptor protein negatively regulates Wnt signaling in intrahepatic cholangiocarcinoma. *Oncol. Lett.* **17**, 2745–2753 (2019).

139. Goeppert, B. *et al.* Global alterations of DNA methylation in cholangiocarcinoma target the Wnt signaling pathway. *Hepatology* **59**, 544–554 (2014).

140. Xue, B., Brown, C. J., Dunker, A. K. & Uversky, V. N. Intrinsically disordered regions of p53 family are highly diversified in evolution. *Biochim. Biophys. Acta* **1834**, 725–738 (2013).

141. Wells, M. *et al.* Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain. *Proc. Natl. Acad. Sci.* **105**, 5762 LP – 5767 (2008).

142. Uversky, V. N. *et al.* Pathological unfoldomics of uncontrolled chaos: intrinsically disordered proteins and human diseases. *Chem. Rev.* **114**, 6844–6879 (2014).

143. Uversky, V. N. p53 Proteoforms and Intrinsic Disorder: An Illustration of the Protein Structure-Function Continuum Concept. *Int. J. Mol. Sci.* **17**, 1874 (2016).

144. Meek, D. W. & Anderson, C. W. Posttranslational Modification of p53: Cooperative Integrators of Function. *Cold Spring Harb. Perspect. Biol.* **1**, (2009).

145. Khoury, M. P. & Bourdon, J.-C. p53 Isoforms: An Intracellular Microprocessor? *Genes Cancer* **2**, 453–465 (2011).

146. Gaglia, G., Guan, Y., Shah, J. V & Lahav, G. Activation and control of p53 tetramerization in individual living cells. *Proc. Natl. Acad. Sci.* **110**, 15497 LP – 15501 (2013).

147. Liu, Y. & Kulesz-Martin, M. p53 protein at the hub of cellular DNA damage response pathways through sequence-specific and non-sequence-specific DNA binding. *Carcinogenesis* **22**, 851–860 (2001).

148. Huang, L. C., Clarkin, K. C. & Wahl, G. M. Sensitivity and selectivity of the DNA damage sensor responsible for activating p53-dependent G1 arrest. *Proc. Natl. Acad. Sci.* **93**, 4827 LP – 4832 (1996).

149. Williams, A. B. & Schumacher, B. p53 in the DNA-Damage-Repair Process. *Cold Spring Harb. Perspect. Med.* **6**, 10.1101/cshperspect.a026070 a026070 (2016).

150. Karni-Schmidt, O., Lokshin, M. & Prives, C. The Roles of MDM2 and MDMX in Cancer. *Annu. Rev. Pathol. Mech. Dis.* **11**, 617–644 (2016).

151. Shadfan, M., Lopez-Pajares, V. & Yuan, Z.-M. MDM2 and MDMX: Alone and together in regulation of p53. *Transl. Cancer Res.* **1**, 88–89 (2012).

152. Pant, V. *et al.* The p53-Mdm2 feedback loop protects against DNA damage by inhibiting p53 activity but is dispensable for p53 stability, development, and longevity. *Genes Dev.* **27**, 1857–1867 (2013).

153. Krenning, L., van den Berg, J. & Medema, R. H. Life or Death after a Break: What Determines the Choice? *Mol. Cell* **76**, 346–358 (2019).

154. Lahav, G. *et al.* Dynamics of the p53-Mdm2 feedback loop in individual cells. *Nat. Genet.* **36**, 147–150 (2004).

155. Choi, J. & Donehower, L. A. p53 in embryonic development: maintaining a fine balance. *Cell. Mol. Life Sci.* **55**, 38–47 (1999).

156. Jain, A. K. & Barton, M. C. p53: emerging roles in stem cells, development and beyond. *Development* **145**, dev158360 (2018).

157. Meletis, K. *et al.* p53 suppresses the self-renewal of adult neural stem cells. *Development* **133**, 363 LP – 369 (2006).

158. Lin, T. *et al.* p53 induces differentiation of mouse embryonic stem cells by suppressing Nanog expression. *Nat. Cell Biol.* **7**, 165–171 (2005).

159. Akdemir, K. C. *et al.* Genome-wide profiling reveals stimulus-specific functions of p53 during differentiation and DNA damage of human embryonic stem cells. *Nucleic Acids Res.* **42**, 205–223 (2014).

160. Jiang, P. *et al.* p53 regulates biosynthesis through direct inactivation of glucose-6-phosphate dehydrogenase. *Nat. Cell Biol.* **13**, 310–316 (2011).

161. Wang, X., Zhao, X., Gao, X., Mei, Y. & Wu, M. A new role of p53 in regulating lipid metabolism. *J. Mol. Cell Biol.* **5**, 147–150 (2012).

162. Krstic, J., Galhuber, M., Schulz, T. J., Schupp, M. & Prokesch, A. p53 as a Dichotomous Regulator of Liver Disease: The Dose Makes the Medicine. *Int. J. Mol. Sci.* **19**, 921 (2018).

163. Borude, P. *et al.* Pleiotropic Role of p53 in Injury and Liver Regeneration after Acetaminophen Overdose. *Am. J. Pathol.* **188**, 1406–1418 (2018).

164. Zou, S. *et al.* Mutational landscape of intrahepatic cholangiocarcinoma. *Nat. Commun.* 5696 (2014) doi:10.1038/ncomms6696.

165. Nakamura, H. *et al.* Genomic spectra of biliary tract cancer. *Nat. Genet.* **47**, 1003–10 (2015).

166. Nepal, C. *et al.* Genomic perturbations reveal distinct regulatory networks in intrahepatic cholangiocarcinoma. *Hepatology* **68**, 949–963 (2018).

167. Ong, C. K. *et al.* Exome sequencing of liver fluke–associated cholangiocarcinoma. *Nat. Genet.* **44**, 690–693 (2012).

168. Chan-On, W. *et al.* Exome sequencing identifies distinct mutational patterns in liver fluke-related and non-infection-related bile duct cancers. *Nat. Genet.* **45**, 1474–1478 (2013).

169. Temko, D., Tomlinson, I. P. M., Severini, S., Schuster-Böckler, B. & Graham, T. A. The effects of mutational processes and selection on driver mutations across cancer types. *Nat. Commun.* **9**, 1857 (2018).

170. Boberg, K. M. *et al.* Cholangiocarcinoma in primary sclerosing cholangitis: K-ras mutations and Tp53 dysfunction are implicated in the neoplastic development. *J. Hepatol.* **32**, 374–380 (2000).

171. Rizzi, P. M. *et al.* p53 Protein overexpression in cholangiocarcinoma arising in primary sclerosing cholangitis. *Gut* **38**, 265–268 (1996).

172. Batheja, N. *et al.* Expression of p53 and PCNA in cholangiocarcinoma and primary sclerosing cholangitis. *Mod. Pathol.* **13**, 1265–1268 (2000).

173. Hill, M. A. *et al.* Kras and Tp53 Mutations Cause Cholangiocyte- and Hepatocyte-Derived Cholangiocarcinoma. 4445–4452 (2018) doi:10.1158/0008-5472.CAN-17-1123.

174. O'Dell, M. R. *et al.* Kras G12Dand p53 mutation cause primary intrahepatic cholangiocarcinoma. *Cancer Res.* **72**, 1557–1567 (2012).

175. El Khatib, M. *et al.* Activation of Notch signaling is required for cholangiocarcinoma progression and is enhanced by inactivation of p53 in vivo. *PLoS One* **8**, 1–9 (2013).

176. Tschaharganeh, D. F. *et al.* p53-Dependent Nestin Regulation Links Tumor Suppression to Cellular Plasticity in Liver Cancer. *Cell* **158**, 579–592 (2014).

177. Liu, F., Yang, X., Geng, M. & Huang, M. Targeting ERK, an Achilles' Heel of the MAPK pathway, in cancer therapy. *Acta Pharm. Sin. B* **8**, 552–562 (2018).

178. Rajalingam, K., Schreck, R., Rapp, U. R. & Albert, Š. Ras oncogenes and their

downstream targets. *Biochim. Biophys. Acta - Mol. Cell Res.* **1773**, 1177–1195 (2007).

179. Sondermann, H. *et al.* Structural Analysis of Autoinhibition in the Ras Activator Son of Sevenless. *Cell* **119**, 393–405 (2004).

180. Bandaru, P., Kondo, Y. & Kuriyan, J. The Interdependent Activation of Son-of-Sevenless and Ras. *Cold Spring Harb. Perspect. Med.* **9**, a031534 (2019).

181. Pamonsinlapatham, P. *et al.* P120-Ras GTPase activating protein (RasGAP): A multi-interacting protein in downstream signaling. *Biochimie* **91**, 320–328 (2009).

182. Cichowski, K. & Jacks, T. NF1 Tumor Suppressor Gene Function: Narrowing the GAP. *Cell* **104**, 593–604 (2001).

183. Mebratu, Y. & Tesfaigzi, Y. How ERK1/2 activation controls cell proliferation and cell death: Is subcellular localization the answer? *Cell Cycle* **8**, 1168–1175 (2009).

184. Cox, A. D. & Der, C. J. Ras history: The saga continues. *Small GTPases* **1**, 2–27 (2010).

185. Hobbs, G. A., Der, C. J. & Rossman, K. L. RAS isoforms and mutations in cancer at a glance. *J. Cell Sci.* **129**, 1287–1292 (2016).

186. Scheffzek, K. *et al.* The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science* **277**, 333–338 (1997).

187. Waters, A. M. & Der, C. J. KRAS: The Critical Driver and Therapeutic Target for Pancreatic Cancer. *Cold Spring Harb. Perspect. Med.* **8**, a031435 (2018).

188. Fan, Z. *et al.* Critical role of KRAS mutation in pancreatic ductal adenocarcinoma. *Transl. Cancer Res. Vol 7, No 6 (December 2018) Transl. Cancer Res.* (2018).

189. Endhardt, K. *et al.* Harvey ras (HRAS) mutations in head and neck cancer (HNC) and dependence on PI3K signaling and resistance to EGFR inhibition. *J. Clin. Oncol.* **32**, 6034 (2014).

190. Malone, E. & Siu, L. L. Precision Medicine in Head and Neck Cancer: Myth or Reality? *Clin. Med. Insights. Oncol.* **12**, 1179554918779581–1179554918779581 (2018).

191. Haigis, K. M. *et al.* Differential effects of oncogenic K-Ras and N-Ras on proliferation, differentiation and tumor progression in the colon. *Nat. Genet.* **40**, 600–608 (2008).

192. To, M. D. *et al.* Kras regulatory elements and exon 4A determine mutation specificity in lung cancer. *Nat. Genet.* **40**, 1240–1244 (2008).

193. Chung, S. I. *et al.* Comparison of liver oncogenic potential among human RAS isoforms. *Oncotarget* **7**, 7354–7366 (2016).

194. Lu, S., Jang, H., Nussinov, R. & Zhang, J. The Structural Basis of Oncogenic Mutations G12, G13 and Q61 in Small GTPase K-Ras4B. *Sci. Rep.* **6**, 21949 (2016).

195. Miller, M. S. & Miller, L. D. RAS Mutations and Oncogenesis: Not all RAS

Mutations are Created Equally. *Front. Genet.* **2**, 100 (2012).

196. Schulien, I. *et al.* The transcription factor c-Jun/AP-1 promotes liver fibrosis during non-alcoholic steatohepatitis by regulating Osteopontin expression. *Cell Death Differ.* **26**, 1688–1699 (2019).

197. Leung, T.-M., Wang, X., Kitamura, N., Fiel, M. I. & Nieto, N. Osteopontin delays resolution of liver fibrosis. *Lab. Invest.* **93**, 1082–1089 (2013).

198. Wen, Y., Jeong, S., Xia, Q. & Kong, X. Role of Osteopontin in Liver Diseases. *Int. J. Biol. Sci.* **12**, 1121–1128 (2016).

199. Wen, Y. *et al.* Defective Initiation of Liver Regeneration in Osteopontin-Deficient Mice after Partial Hepatectomy due to Insufficient Activation of IL-6/Stat3 Pathway. *Int. J. Biol. Sci.* **11**, 1236–1247 (2015).

200. Zhao, H. *et al.* The role of osteopontin in the progression of solid organ tumour. *Cell Death Dis.* **9**, 356 (2018).

201. Zheng, Y. *et al.* Osteopontin promotes metastasis of intrahepatic cholangiocarcinoma through recruiting MAPK1 and mediating Ser675 phosphorylation of β-Catenin. *Cell Death Dis.* **9**, (2018).

202. Kim, M. *et al.* Merlin inhibits Wnt/β-catenin signaling by blocking LRP6 phosphorylation. *Cell Death Differ.* **23**, 1638–1647 (2016).

203. Konsavage Jr, W. M., Kyler, S. L., Rennoll, S. A., Jin, G. & Yochum, G. S. Wnt/β-catenin signaling regulates Yes-associated protein (YAP) gene expression in colorectal carcinoma cells. *J. Biol. Chem.* **287**, 11730–11739 (2012).

204. Jusakul, A. *et al. Whole-genome and epigenomic landscapes of etiologically distinct subtypes of cholangiocarcinoma. Cancer Discovery* vol. 7 (2017).

205. Ma, L.-G. *et al.* LKB1 inhibits the proliferation of gastric cancer cells by suppressing the nuclear translocation of Yap and β-catenin. *Int. J. Mol. Med.* **37**, 1039–1048 (2016).

206. Mungamuri, S. K., Yang, X., Thor, A. D. & Somasundaram, K. Survival signaling by Notch1: mammalian target of rapamycin (mTOR)-dependent inhibition of p53. *Cancer Res.* **66**, 4715–4724 (2006).

207. Beverly, L. J., Felsher, D. W. & Capobianco, A. J. Suppression of p53 by Notch in lymphomagenesis: implications for initiation and regression. *Cancer Res.* **65**, 7159–7168 (2005).

208. Colaluca, I. N. *et al.* A Numb-Mdm2 fuzzy complex reveals an isoform-specific involvement of Numb in breast cancer. *J. Cell Biol.* **217**, 745–762 (2018).

209. Farshidfar, F. *et al.* Integrative Genomic Analysis of Cholangiocarcinoma Identifies Distinct IDH-Mutant Molecular Profiles. *Cell Rep.* **18**, 2780–2794 (2017).

210. Gu, X. & Manautou, J. E. Molecular mechanisms underlying chemical liver injury. *Expert Rev. Mol. Med.* **14**, e4–e4 (2012).

211. Goldkind, L. & Laine, L. A systematic review of NSAIDs withdrawn from the

market due to hepatotoxicity: lessons learned from the bromfenac experience. *Pharmacoepidemiol. Drug Saf.* **15**, 213–220 (2006).

212. Kensler, T. W., Roebuck, B. D., Wogan, G. N. & Groopman, J. D. Aflatoxin: a 50-year odyssey of mechanistic and translational toxicology. *Toxicol. Sci.* **120 Suppl**, S28–S48 (2011).

213. Letouzé, E. *et al.* Mutational signatures reveal the dynamic interplay of risk factors and cellular processes during liver tumorigenesis. *Nat. Commun.* **8**, 1315 (2017).

214. Huang, M. N. *et al.* Genome-scale mutational signatures of aflatoxin in cells, mice, and human tumors. *Genome Res.* **27**, 1475–1486 (2017).

215. Sukowati, C. H. C. *et al.* Significance of hepatitis virus infection in the oncogenic initiation of hepatocellular carcinoma. *World J. Gastroenterol.* **22**, 1497–1512 (2016).

216. Gurtsevitch, V. E. Human oncogenic viruses: hepatitis B and hepatitis C viruses and their role in hepatocarcinogenesis. *Biochemistry. (Mosc).* **73**, 504–513 (2008).

217. Fujimoto, A. *et al.* Whole-genome mutational landscape of liver cancers displaying biliary phenotype reveals hepatitis impact and molecular diversity. *Nat. Commun.* **6**, 6120 (2015).

218. Fried, B., Reddy, A. & Mayer, D. Helminths in human carcinogenesis. *Cancer Lett.* **305**, 239–249 (2011).

219. Chaiyadet, S. *et al.* Carcinogenic Liver Fluke Secretes Extracellular Vesicles That Promote Cholangiocytes to Adopt a Tumorigenic Phenotype. *J. Infect. Dis.* **212**, 1636–1645 (2015).

220. Sripa, B. *et al.* The tumorigenic liver fluke Opisthorchis viverrini - multiple pathways to cancer. *Trends Parasitol.* **28**, 395–407 (2012).

221. Vennervald, B. J. & Polman, K. Helminths and malignancy. *Parasite Immunol.* **31**, 686–696 (2009).

222. Sia, D. *et al.* Massive parallel sequencing uncovers actionable FGFR2–PPHLN1 fusion and ARAF mutations in intrahepatic cholangiocarcinoma. *Nat. Commun.* **6**, 6087 (2015).

223. Sia, D. *et al.* Integrative molecular analysis of intrahepatic cholangiocarcinoma reveals 2 classes that have different outcomes. *Gastroenterology* **144**, 829–840 (2013).

224. Dieci, M. V., Arnedos, M., Andre, F. & Soria, J. C. Fibroblast Growth Factor Receptor Inhibitors as a Cancer Treatment: From a Biologic Rationale to Medical Perspectives. *Cancer Discov.* **3**, 264 LP – 279 (2013).

225. Arai, Y. *et al.* Fibroblast growth factor receptor 2 tyrosine kinase fusions define a unique molecular subtype of cholangiocarcinoma. *Hepatology* **59**, 1427–1434 (2014).

226. Wang, Y. *et al.* Antitumor effect of FGFR inhibitors on a novel

cholangiocarcinoma patient derived xenograft mouse model endogenously expressing an FGFR2-CCDC6 fusion protein. *Cancer Lett.* **380**, 163–173 (2016).

227. Dang, L. & Su, S. M. Isocitrate Dehydrogenase Mutation and (R) -2-Hydroxyglutarate : From Basic Discovery to Therapeutics Development. 1–27 (2017) doi:10.1146/annurev-biochem-061516-044732.

228. Dang, L. *et al.* Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature* **462**, 739–744 (2009).

229. Turcan, S. *et al.* IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. *Nature* **483**, 479–83 (2012).

230. Lu, C., Ward, P., Kapoor, G. & Rohle, D. IDH mutation impairs histone demethylation and results in a block to cell differentiation. *Nature* **483**, 474–478 (2012).

231. Saha, S. K., Parachoniak, C. A. & Bardeesy, N. IDH mutations in liver cell plasticity and biliary cancer. *Cell Cycle* **13**, 3176–3182 (2014).

232. Rizvi, S. & Borad, M. J. The rise of the FGFR inhibitor in advanced biliary cancer: The next cover of time magazine? *J. Gastrointest. Oncol.* **7**, 789–796 (2016).

233. Rizvi, S., Khan, S. A., Hallemeier, C. L., Kelley, R. K. & Gores, G. J. Cholangiocarcinoma-evolving concepts and therapeutic strategies. *Nat. Rev. Clin. Oncol.* **15**, 95–111 (2018).

234. Finn, R. S. *et al.* Phase 1b investigation of the MEK inhibitor binimetinib in patients with advanced or metastatic biliary tract cancer. *Invest. New Drugs* **36**, 1037–1043 (2018).

235. Bekaii-Saab, T. *et al.* Multi-Institutional Phase II Study of Selumetinib in Patients With Metastatic Biliary Cancers. *J. Clin. Oncol.* **29**, 2357–2363 (2011).

236. Dong, M. *et al.* Efficacy of MEK inhibition in a K-Ras-driven cholangiocarcinoma preclinical model. *Cell Death Dis.* **9**, 31 (2018).

237. Lowery, M. A. *et al.* Safety and activity of ivosidenib in patients with IDH1-mutant advanced cholangiocarcinoma: a phase 1 study. *Lancet Gastroenterol. Hepatol.* **4**, 711–720 (2019).

238. Javle, M. M. *et al.* Infigratinib versus gemcitabine plus cisplatin multicenter, open-label, randomized, phase 3 study in patients with advanced cholangiocarcinoma with FGFR2 gene fusions/translocations: The PROOF trial. *J. Clin. Oncol.* **37**, TPS4155–TPS4155 (2019).

239. Javle, M. *et al.* LBA28Updated results from a phase II study of infigratinib (BGJ398), a selective pan-FGFR kinase inhibitor, in patients with previously treated advanced cholangiocarcinoma containing FGFR2 fusions. *Ann. Oncol.* **29**, (2018).

240. Schneider, G., Schmidt-supprian, M., Rad, R. & Saur, D. Tissue-specific tumorigenesis: context matters Günter. doi:10.1038/nrc.2017.5.

241. Morris 4th, J. P., Wang, S. C. & Hebrok, M. KRAS, Hedgehog, Wnt and the twisted

developmental biology of pancreatic ductal adenocarcinoma. *Nat. Rev. Cancer* **10**, 683–695 (2010).

242. Kang, T. *et al.* hepatocytes limits liver cancer development. *Nature* **479**, 547–551 (2011).

243. Kang, T.-W. *et al.* Senescence surveillance of pre-malignant hepatocytes limits liver cancer development. *Nature* **479**, 547–551 (2011).

244. Nguyen, A. T. *et al.* A high level of liver-specific expression of oncogenic KrasV12 drives robust liver tumorigenesis in transgenic zebrafish. *Dis. Model. & Mech.* **4**, 801 LP – 813 (2011).

245. Tabibian, J. H., O'Hara, S. P., Splinter, P. L., Trussoni, C. E. & Larusso, N. F. Cholangiocyte senescence by way of N-Ras activation is a characteristic of primary sclerosing cholangitis. *Hepatology* **59**, 2263–2275 (2014).

246. O'Hara, S. P. *et al.* Cholangiocyte N-Ras protein mediates lipopolysaccharide-induced interleukin 6 secretion and proliferation. *J. Biol. Chem.* **286**, 30352–30360 (2011).

247. Isomoto, H. *et al.* Sustained IL-6/STAT-3 signaling in cholangiocarcinoma cells due to SOCS-3 epigenetic silencing. *Gastroenterology* **132**, 384–396 (2007).

248. Labib, P. L., Goodchild, G. & Pereira, S. P. Molecular Pathogenesis of Cholangiocarcinoma. *BMC Cancer* **19**, 1–16 (2019).

249. Braconi, C., Huang, N. & Patel, T. MicroRNA-dependent regulation of DNA methyltransferase-1 and tumor suppressor gene expression by interleukin-6 in human malignant cholangiocytes. *Hepatology* **51**, 881–890 (2010).

250. Frampton, G. *et al.* Interleukin-6-driven progranulin expression increases cholangiocarcinoma growth by an Akt-dependent mechanism. *Gut* **61**, 268–277 (2012).

251. Smout, M. J. *et al.* A granulin-like growth factor secreted by the carcinogenic liver fluke, Opisthorchis viverrini, promotes proliferation of host cells. *PLoS Pathog.* **5**, e1000611–e1000611 (2009).

252. Young, N. D. *et al.* The Opisthorchis viverrini genome provides insights into life in the bile duct. *Nat. Commun.* **5**, 1–11 (2014).

253. Manchado, E. *et al.* A combinatorial strategy for treating KRAS-mutant lung cancer. *Nature* **534**, 647–651 (2016).

254. Brandt, R. *et al.* Cell type-dependent differential activation of ERK by oncogenic KRAS in colon cancer and intestinal epithelium. *Nat. Commun.* **10**, 2919 (2019).

255. Yuan, T. L. *et al.* Differential Effector Engagement by Oncogenic KRAS. *Cell Rep.* **22**, 1889–1902 (2018).

256. Cardinale, V. Multiple cells of origin in cholangiocarcinoma underlie biological, epidemiological and clinical heterogeneity. *World J. Gastrointest. Oncol.* **4**, 94 (2012).

257. Polak, P. *et al.* Cell-of-origin chromatin organization shapes the mutational landscape of cancer. *Nature* **518**, 360–364 (2015).

258. Hoadley, K. A. *et al.* Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. *Cell* **173**, 291-304.e6 (2018).

259. Salvadores, M., Mas-Ponte, D. & Supek, F. Passenger mutations accurately classify human tumors. *PLOS Comput. Biol.* **15**, e1006953 (2019).

260. Curtius, K., Wright, N. A. & Graham, T. A. An evolutionary perspective on field cancerization. *Nat. Rev. Cancer* **18**, 19–32 (2017).

261. Lipinski, K. A. *et al.* Cancer Evolution and the Limits of Predictability in Precision Cancer Medicine. *Trends in Cancer* vol. 2 49–63 (2016).

262. Phillips, P. C. Epistasis - The essential role of gene interactions in the structure and evolution of genetic systems. *Nat. Rev. Genet.* **9**, 855–867 (2008).

263. McGranahan, N. & Swanton, C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. *Cell* **168**, 613–628 (2017).

264. van de Haar, J. *et al.* Identifying Epistasis in Cancer Genomes: A Delicate Affair. *Cell* **177**, 1375–1383 (2019).

265. Marusyk, A., Almendro, V. & Polyak, K. Intra-tumour heterogeneity: a looking glass for cancer? *Nat. Rev. Cancer* **12**, 323–334 (2012).

266. Bauer, B., Siebert, R. & Traulsen, A. Cancer initiation with epistatic interactions between driver and passenger mutations. *J. Theor. Biol.* **358**, 52–60 (2014).

267. Wodarz, D., Newell, A. C. & Komarova, N. L. Passenger mutations can accelerate tumour suppressor gene inactivation in cancer evolution. *J. R. Soc. Interface* **15**, 20170967 (2018).

268. Gonzalez-Perez, A. *et al.* Computational approaches to identify functional genetic variants in cancer genomes. *Nat. Methods* **10**, 723–729 (2013).

269. Ding, L., Wendl, M. C., McMichael, J. F. & Raphael, B. J. Expanding the computational toolbox for mining cancer genomes. *Nat. Rev. Genet.* **15**, 556–570 (2014).

270. Adzhubei, I., Jordan, D. M. & Sunyaev, S. R. *Predicting functional effect of human missense mutations using PolyPhen-2*. *Current Protocols in Human Genetics* vol. 2 (2013).

271. Kumar, P., Henikoff, S. & Ng, P. C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* **4**, 1073–1081 (2009).

272. Reva, B., Antipin, Y. & Sander, C. Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res.* **39**, e118–e118 (2011).

273. Reimand, J. & Bader, G. D. Systematic analysis of somatic mutations in phosphorylation signaling predicts novel cancer drivers. *Mol. Syst. Biol.* **9**, 637

(2013).

274. Reimand, J., Wagih, O. & Bader, G. D. The mutational landscape of phosphorylation signaling in cancer. *Sci. Rep.* **3**, 2651 (2013).

275. Gonzalez-Perez, A. & Lopez-Bigas, N. Functional impact bias reveals cancer drivers. *Nucleic Acids Res.* **40**, 1–10 (2012).

276. Tamborero, D., Gonzalez-Perez, A. & Lopez-Bigas, N. OncodriveCLUST: Exploiting the positional clustering of somatic mutations to identify cancer genes. *Bioinformatics* **29**, 2238–2244 (2013).

277. Kamburov, A. *et al.* Comprehensive assessment of cancer missense mutation clustering in protein structures. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E5486–E5495 (2015).

278. Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013).

279. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29 (2000).

280. Leiserson, M. D. M. *et al.* Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes. *Nat. Genet.* **47**, 106–114 (2014).

281. Shannon, P. *et al.* Cytoscape: A Software Environment for Integrated Models. **13**, 426 (1971).

282. Fabregat, A. *et al.* The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* **46**, D649–D655 (2017).

283. Tokheim, C. J., Papadopoulos, N., Kinzler, K. W., Vogelstein, B. & Karchin, R. Evaluating the evaluation of cancer driver genes. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 14330–14335 (2016).

284. Coveney, P. V., Dougherty, E. R. & Highfeld, R. R. Big data need big theory too. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **374**, (2016).

285. Nussinov, R., Jang, H., Tsai, C. J. & Cheng, F. *Precision medicine and driver mutations: Computational methods, functional assays and conformational principles for interpreting cancer drivers*. *PLoS Computational Biology* vol. 15 (2019).

286. Weber, J. *et al.* CRISPR/Cas9 somatic multiplex-mutagenesis for high-throughput functional cancer genomics in mice. *Proc. Natl. Acad. Sci.* **112**, 13982–13987 (2015).

287. Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* **11**, 783–784 (2014).

288. Shalem, O. *et al.* Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. *Science (80-.).* **343**, 84 LP – 87 (2014).

289. Read, A., Gao, S., Batchelor, E. & Luo, J. Flexible CRISPR library construction using

parallel oligonucleotide retrieval. *Nucleic Acids Res.* **45**, e101–e101 (2017).

290. Li, W. *et al.* Quality control, modeling, and visualization of CRISPR screens with MAGeCK-VISPR. *Genome Biol.* **16**, 1–13 (2015).

291. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. *Methods* **25**, 402–408 (2001).

292. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

293. DePristo, M. A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).

294. Garrison, E. & Marth, G. Haplotype-based variant detection from short-read sequencing. 1–9 (2012).

295. Lai, Z. *et al.* VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. *Nucleic Acids Res.* **44**, e108–e108 (2016).

296. Koboldt, D. C. *et al.* VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics* **25**, 2283–2285 (2009).

297. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016).

298. Gonzalez-Perez, A. *et al.* IntOGen-mutations identifies cancer drivers across tumor types. *Nat. Methods* **10**, 1081–1082 (2013).

299. Gonzalez-Perez, A., Deu-Pons, J. & Lopez-Bigas, N. Improving the prediction of the functional impact of cancer mutations by baseline tolerance transformation. *Genome Med.* **4**, 89 (2012).

300. Lawrence, M. S. *et al.* Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–8 (2013).

301. Rosenthal, R., McGranahan, N., Herrero, J., Taylor, B. S. & Swanton, C. deconstructSigs: Delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. *Genome Biol.* **17**, 1–11 (2016).

302. Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415–21 (2013).

303. Wu, G., Feng, X. & Stein, L. Research A human functional protein interaction network and its application to cancer data analysis. *Genome Biol.* **11**, (2012).

304. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2012).

305. Rausch, T. *et al.* DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics* **28**, i333–i339 (2012).

306. Reimand, J., Kull, M., Peterson, H., Hansen, J. & Vilo, J. g:Profiler--a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic*

Acids Res. **35**, W193–W200 (2007).

307. The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* **47**, D330–D338 (2018).

308. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44**, D457–D462 (2015).

309. Matys, V. *et al.* TRANSFAC® and its module TRANSCompel®: transcriptional gene regulation in eukaryotes. *Nucleic Acids Res.* **34**, D108–D110 (2006).

310. Chou, C.-H. *et al.* miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res.* **46**, D296–D302 (2017).

311. Wickham, H. *Elegant Graphics for Data Analysis: ggplot2*. *Applied Spatial Data Analysis with R* (2008). doi:10.1007/978-0-387-78171-6.

312. Griffith, M. *et al.* Optimizing Cancer Genome Sequencing and Analysis. *Cell Syst.* **1**, 210–223 (2015).

313. Kikuno, R. HUGE: a database for human KIAA proteins, a 2004 update integrating HUGEppi and ROUGE. *Nucleic Acids Res.* **32**, 502D – 504 (2004).

314. Tate, J. G. *et al.* COSMIC: The Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res.* **47**, D941–D947 (2019).

315. Sondka, Z. *et al.* The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. *Nat. Rev. Cancer* **18**, 696–705 (2018).

316. Jiang, W. *et al.* The PIK3CA E542K and E545K mutations promote glycolysis and proliferation via induction of the β-catenin/SIRT3 signaling pathway in cervical cancer. *J. Hematol. Oncol.* **11**, 1–15 (2018).

317. Li, J. *et al.* SETD2: An epigenetic modifier with tumor suppressor functionality. *Oncotarget* **7**, 50719–50734 (2016).

318. Bihr, S. *et al.* Expression and Mutation Patterns of PBRM1, BAP1 and SETD2 Mirror Specific Evolutionary Subtypes in Clear Cell Renal Cell Carcinoma. *Neoplasia (United States)* **21**, 247–256 (2019).

319. Newman, M. E. J. Modularity and Community Structure in Networks. *Proc. Natl. Acad. Sci.* (2006).

320. Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L. A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* **31**, 233–239 (2013).

321. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science (80-.).* **339**, 823–826 (2013).

322. Jinek, M. *et al.* A Programmable Dual-RNA – Guided. **337**, 816–822 (2012).

323. Jiang, F. & Doudna, J. A. CRISPR – Cas9 Structures and Mechanisms. *Annu.Rev.Biophys* **46**, 505–529 (2017).

324. Aird, E. J., Lovendahl, K. N., St. Martin, A., Harris, R. S. & Gordon, W. R. Increasing

Cas9-mediated homology-directed repair efficiency through covalent tethering of DNA repair template. *Commun. Biol.* **1**, (2018).

325. Lundstrom, K. Viral Vectors in Gene Therapy. *Dis. (Basel, Switzerland)* **6**, 42 (2018).

326. Tros de Ilarduya, C., Sun, Y. & Düzgüneş, N. Gene delivery by lipoplexes and polyplexes. *Eur. J. Pharm. Sci.* **40**, 159–170 (2010).

327. Sebestyén, M. G. *et al.* Mechanism of plasmid delivery by hydrodynamic tail vein injection. I. Hepatocyte uptake of various molecules. *J. Gene Med.* **8**, 852–873 (2006).

328. Kanefuji, T. *et al.* Hemodynamics of a hydrodynamic injection. *Mol. Ther. - Methods Clin. Dev.* **1**, 14029 (2014).

329. Zhang, G. *et al.* Hydroporation as the mechanism of hydrodynamic delivery. *Gene Ther.* **11**, 675–682 (2004).

330. Muzumdar, M. D., Tasic, B., Miyamichi, K., Li, L. & Luo, L. A Global Double-Fluorescent Cre Reporter Mouse. *Genesis* **45**, 593–605 (2007).

331. Du, W., Hu, J. K. H., Du, W. & Klein, O. D. Lineage tracing of epithelial cells in developing teeth reveals two strategies for building signaling centers. *J. Biol. Chem.* **292**, 15062–15069 (2017).

332. Snyder, C. S. *et al.* A dual-color genetically engineered mouse model for multispectral imaging of the pancreatic microenvironment. *Pancreas* **42**, 952–958 (2013).

333. Alapati, D. *et al.* In utero gene editing for monogenic lung disease. *Sci. Transl. Med.* **11**, 1–14 (2019).

334. Cronin, J., Zhang, X.-Y. & Reiser, J. Altering the tropism of lentiviral vectors through pseudotyping. *Curr. Gene Ther.* **5**, 387–398 (2005).

335. Sohlenius-Sternbeck, A. K. Determination of the hepatocellularity number for human, dog, rabbit, rat and mouse livers from protein concentration measurements. *Toxicol. Vitr.* **20**, 1582–1586 (2006).

336. Nascimento-Sales, M. *et al.* Is the FVB/N mouse strain truly resistant to diet-induced obesity? *Physiol. Rep.* **5**, 1–12 (2017).

337. Zhan, T., Rindtorff, N. & Boutros, M. Wnt signaling in cancer. *Nat. Publ. Gr.* **36**, 1461–1473 (2016).

338. Tward, A. D. *et al.* Distinct pathways of genomic progression to benign and malignant tumors of the liver. *Proc. Natl. Acad. Sci.* **104**, 14771 LP – 14776 (2007).

339. Paige, A. J. W. Redefining tumour suppressor genes: Exceptions to the two-hit hypothesis. *Cell. Mol. Life Sci.* **60**, 2147–2163 (2003).

340. Knudson, a G. Two genetic hits (more or less) to cancer. *Nat. Rev. Cancer* **1**, 157–162 (2001).

341. Ding, L. *et al.* Perspective on Oncogenic Processes at the End of the Beginning of

Cancer Genomics. *Cell* **173**, 305-320.e10 (2018).

342. Wang, J. *et al.* Notch2 controls hepatocyte-derived cholangiocarcinoma formation in mice. *Oncogene* **37**, 3229–3242 (2018).

343. Strazzabosco, M. & Fabris, L. Development of the bile ducts: Essentials for the clinical hepatologist. *J. Hepatol.* **56**, 1159–1170 (2012).

344. Shibuya, K. *et al.* A cluster of 21 keratin-associated protein genes within introns of another gene on human chromosome 21q22.3. *Genomics* **83**, 679–693 (2004).

345. Lingenfelter, P. A. *et al.* Expression and conservation of processed copies of the RBMX gene. *Mamm. Genome* **12**, 538–545 (2001).

346. Cheung, H.-H., Yang, Y., Lee, T.-L., Rennert, O. & Chan, W.-Y. Hypermethylation of genes in testicular embryonal carcinomas. *Br. J. Cancer* **114**, 230 (2015).

347. Friedli, M. & Trono, D. The Developmental Control of Transposable Elements and the Evolution of Higher Species. *Annu. Rev. Cell Dev. Biol.* **31**, 429–451 (2015).

348. Yang, P., Wang, Y. & Macfarlan, T. S. The Role of KRAB-ZFPs in Transposable Element Repression and Mammalian Evolution. *Trends Genet.* **33**, 871–881 (2017).

349. Joung, J. *et al.* Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nat. Protoc.* **12**, 828–863 (2017).

350. O'Hagan, S., Wright Muelas, M., Day, P. J., Lundberg, E. & Kell, D. B. GeneGini: Assessment via the Gini Coefficient of Reference 'Housekeeping' Genes and Diverse Human Transporter Expression Profiles. *Cell Syst.* **6**, 230-244.e1 (2018).

351. Wittebolle, L. *et al.* Initial community evenness favours functionality under selective stress. *Nature* **458**, 623–626 (2009).

352. Wei, M. *et al.* Multiple cellular origins and molecular evolution of intrahepatic cholangiocarcinoma. *Cancer Lett.* **379**, 253–261 (2016).

353. Charlesworth, C. T. *et al.* Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nat. Med.* **25**, 249–254 (2019).

354. Wagner, D. L. *et al.* High prevalence of Streptococcus pyogenes Cas9-reactive T cells within the adult human population. *Nat. Med.* **25**, 242–248 (2019).

355. Crudele, J. M. & Chamberlain, J. S. Cas9 immunity creates challenges for CRISPR gene editing therapies. *Nat. Commun.* **9**, 3497 (2018).

356. Wang, D. *et al.* Adenovirus-Mediated Somatic Genome Editing of Pten by CRISPR/Cas9 in Mouse Liver in Spite of Cas9-Specific Immune Responses. *Hum. Gene Ther.* **26**, 432–442 (2015).

357. Ajina, R. *et al.* SpCas9-expression by tumor cells can cause T cell-dependent tumor rejection in immunocompetent mice. *Oncoimmunology* **8**, e1577127–e1577127 (2019).

358. Liau, J.-Y. *et al.* Morphological subclassification of intrahepatic cholangiocarcinoma: etiological, clinicopathological, and molecular features.

Mod. Pathol. **27**, 1163–73 (2014).

359. Vijgen, S., Terris, B. & Rubbia-Brandt, L. Pathology of intrahepatic cholangiocarcinoma. *HepatoBiliary Surg. Nutr.* **6**, 22–34 (2017).

360. Jain, R., Fischer, S., Serra, S. & Chetty, R. The Use of Cytokeratin 19 (CK19) Immunohistochemistry in Lesions of the Pancreas, Gastrointestinal Tract, and Liver. *Appl. Immunohistochem. Mol. Morphol.* **18**, (2010).

361. Lau, S. K., Prakash, S., Geller, S. A. & Alsabeh, R. Comparative immunohistochemical profile of hepatocellular carcinoma, cholangiocarcinoma, and metastatic adenocarcinoma. *Hum. Pathol.* **33**, 1175–1181 (2002).

362. Nakajima, T. *et al.* Intrahepatic cholangiocarcinoma with sarcomatous change. Clinicopathologic and immunohistochemical evaluation of seven cases. *Cancer* **72**, 1872–1877 (1993).

363. Aishima, S. *et al.* Prognostic impact of cholangiocellular and sarcomatous components in combined hepatocellular and cholangiocarcinoma. *Hum. Pathol.* **37**, 283–291 (2006).

364. Carpino, G. *et al.* Matrisome analysis of intrahepatic cholangiocarcinoma unveils a peculiar cancer - associated extracellular matrix structure. *Clin. Proteomics* 1–12 (2019) doi:10.1186/s12014-019-9257-x.

365. Sirica, A. E. The role of cancer-associated myofibroblasts in intrahepatic cholangiocarcinoma. *Nat. Rev. Gastroenterol. Hepatol.* **9**, 44–54 (2012).

366. Xu, S. *et al.* The role of collagen in cancer: from bench to bedside. *J. Transl. Med.* **17**, 1–22 (2019).

367. Prakobwong, S. *et al.* Plasma hydroxyproline, MMP-7 and collagen I as novel predictive risk markers of hepatobiliary disease-associated cholangiocarcinoma. *Int. J. Cancer* **131**, E416–E424 (2012).

368. Erdogan, B. *et al.* Cancer-associated fibroblasts promote directional cancer cell migration by aligning fibronectin. *J. Cell Biol.* **216**, 3799 LP – 3816 (2017).

369. Sun, S., Genovese, F. & Karsdal, M. A. *Type VI collagen*. *Biochemistry of Collagens, Laminins and Elastin* (Elsevier Inc., 2019). doi:10.1016/b978-0-12-817068-7.00006-9.

370. Iyengar, P. *et al.* Adipocyte-derived collagen VI affects early mammary tumor progression in vivo, demonstrating a critical interaction in the tumor/stroma microenvironment. *J. Clin. Invest.* **115**, 1163–1176 (2005).

371. Park, J. & Scherer, P. E. Adipocyte-derived endotrophin promotes malignant tumor progression. *J. Clin. Invest.* **122**, 4243–4256 (2012).

372. Cescon, M., Gattazzo, F., Chen, P. & Bonaldo, P. Collagen VI at a glance. *J. Cell Sci.* **128**, 3525–3531 (2015).

373. Chen, P., Cescon, M. & Bonaldo, P. Collagen VI in cancer and its biological mechanisms. *Trends Mol. Med.* **19**, 410–417 (2013).

374. Sherman-Baust, C. A. *et al.* Remodeling of the extracellular matrix through overexpression of collagen VI contributes to cisplatin resistance in ovarian cancer cells. *Cancer Cell* **3**, 377–386 (2003).

375. Park, J., Morley, T. S. & Scherer, P. E. Inhibition of endotrophin, a cleavage product of collagen VI, confers cisplatin sensitivity to tumours. *EMBO Mol. Med.* **5**, 935–948 (2013).

376. Torbenson, M. *et al.* Hepatocellular carcinomas show abnormal expression of fibronectin protein. *Mod. Pathol.* **15**, 826–830 (2002).

377. Soejima, Y. *et al.* Integrins αvβ6, α6β4 and α3β1 are down-regulated in cholangiolocellular carcinoma but not cholangiocarcinoma. *Hepatol. Res.* **44**, E320–E334 (2014).

378. Ryu, S., Jimi, S., Eura, Y., Kato, T. & Takebayashi, S. Retention of intracellular fibronectin expression in primary and metastatic thyroid carcinoma: an immunohistochemical study. *Cancer Lett.* **133**, 215–222 (1998).

379. Tsuchida, T. & Friedman, S. L. Mechanisms of hepatic stellate cell activation. *Nat. Rev. Gastroenterol. Hepatol.* **14**, 397–411 (2017).

380. Marcher, A.-B. *et al.* Transcriptional regulation of Hepatic Stellate Cell activation in NASH. *Sci. Rep.* **9**, 2324 (2019).

381. Kraft, K. *et al.* Deletions, Inversions, Duplications: Engineering of Structural Variants using CRISPR/Cas in Mice. *Cell Rep.* **10**, 833–839 (2015).

382. Martínez-Iglesias, O. A. *et al.* Autoregulatory loop of nuclear corepressor 1 expression controls invasion, tumor growth, and metastasis. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E328–E337 (2016).

383. Ma, G., Dai, W., Sang, A., Yang, X. & Li, Q. Roles of ZIC family genes in human gastric cancer. *Int. J. Mol. Med.* **38**, 259–266 (2016).

384. Wang, L. J. *et al.* ZIC1 is downregulated through promoter hypermethylation in gastric cancer. *Biochem. Biophys. Res. Commun.* **379**, 959–963 (2009).

385. Gan, L. *et al.* ZIC1 Is Downregulated through Promoter Hypermethylation, and Functions as a Tumor Suppressor Gene in Colorectal Cancer. *PLoS One* **6**, e16916 (2011).

386. Zhong, J. *et al.* ZIC1 modulates cell-cycle distributions and cell migration through regulation of sonic hedgehog, PI3K and MAPK signaling pathways in gastric cancer. *BMC Cancer* **12**, 290 (2012).

387. Yeh, M., Oh, C. S., Yoo, J. Y., Kaur, B. & Lee, T. J. Pivotal role of microRNA-138 in human cancers. *Am. J. Cancer Res.* **9**, 1118–1126 (2019).

388. Davis, J. A., Saunders, S. J., Mann, M. & Backofen, R. Combinatorial ensemble miRNA target prediction of co-regulation networks with non-prediction data. *Nucleic Acids Res.* **45**, 8745–8757 (2017).

389. Han, W. *et al.* ZIC1 acts a tumor suppressor in breast cancer by targeting survivin. *Int. J. Oncol.* **53**, 937–948 (2018).

390. Yang, R. *et al.* miR-138-5p contributes to cell proliferation and invasion by targeting Survivin in bladder cancer cells. *Mol. Cancer* **15**, 82 (2016).

391. Chang, Q. *et al.* Survivin expression induced by doxorubicin in cholangiocarcinoma. *World J. Gastroenterol.* **10**, 415–418 (2004).

392. Javle, M. M. *et al.* Nuclear survivin expression predicts poor outcome in cholangiocarcinoma. *Hepatogastroenterology.* **51**, 1653–1657 (2004).

393. Ou, B. *et al.* Plk2 promotes tumor growth and inhibits apoptosis by targeting Fbxw7/Cyclin E in colorectal cancer. *Cancer Lett.* **380**, 457–466 (2016).

394. Fingas, C. D. *et al.* Polo-like kinase 2 is a mediator of hedgehog survival signaling in cholangiocarcinoma. *Hepatology* **58**, 1362–1374 (2013).

395. Syed, N. *et al.* Polo-like Kinase Plk2 Is an Epigenetic Determinant of Chemosensitivity and Clinical Outcomes in Ovarian Cancer. *Cancer Res.* **71**, 3317 LP – 3327 (2011).

396. Lee, J. H., Kim, M. S., Yoo, N. J. & Lee, S. H. Frameshift mutation and loss of expression of PLK2, a serine/threonine kinase-encoding gene, in colorectal cancers. *Pathol. - Res. Pract.* **213**, 1019–1020 (2017).

397. Liu, L. Y. *et al.* Silencing of polo-like kinase 2 increases cell proliferation and decreases apoptosis in SGC-7901 gastric cancer cells. *Mol. Med. Rep.* **11**, 3033–3038 (2015).

398. Matthew, E. M. *et al.* Plk2 Loss Commonly Occurs in Colorectal Carcinomas but not Adenomas: Relationship to mTOR Signaling. *Neoplasia* **20**, 244–255 (2018).

399. Pellegrino, R. *et al.* Oncogenic and tumor suppressive roles of polo-like kinases in human hepatocellular carcinoma. *Hepatology* **51**, 857–868 (2010).

400. Tran, T. S., Kolodkin, A. L. & Bharadwaj, R. Semaphorin Regulation of Cellular Morphology. *Annu. Rev. Cell Dev. Biol.* **23**, 263–292 (2007).

401. Takamatsu, H. & Kumanogoh, A. Diverse roles for semaphorin plexin signaling in the immune system. *Trends Immunol.* **33**, 127–135 (2012).

402. Tong, Y. *et al.* Structure and Function of the Intracellular Region of the Plexin-B1 Transmembrane Receptor. *J. Biol. Chem.* **284**, 35962–35972 (2009).

403. Liu, H. & Zhao, H. Prognosis related miRNAs, DNA methylation, and epigenetic interactions in lung adenocarcinoma. *Neoplasma* **66**, 487–493 (2019).

404. Ambrogio, C. *et al.* KRAS-driven lung adenocarcinoma: combined DDR1/Notch inhibition as an effective therapy. *ESMO Open* **1**, e000076 (2016).

405. Rittinger, K. & Ikeda, F. Linear ubiquitin chains: Enzymes, mechanisms and biology. *Open Biol.* **7**, (2017).

406. Joo, D. *et al.* Regulation of Linear Ubiquitin Chain Assembly Complex by Caspase-Mediated Cleavage of RNF31. *Mol. Cell. Biol.* **36**, 3010–3018 (2016).

407. Shimizu, Y. *et al.* The Linear ubiquitin chain assembly complex acts as a liver tumor suppressor and inhibits hepatocyte apoptosis and hepatitis. *Hepatology*

65, 1963–1978 (2017).

408. Schmidt, C., Sciacovelli, M. & Frezza, C. Fumarate hydratase in cancer: A multifaceted tumour suppressor. *Semin. Cell Dev. Biol.* (2019) doi:https://doi.org/10.1016/j.semcdb.2019.05.002.

409. Pascoe, H. G., Wang, Y. & Zhang, X. Structural mechanisms of plexin signaling. *Prog. Biophys. Mol. Biol.* **118**, 161–168 (2015).

410. Wylie, T., Garg, R., Ridley, A. J. & Conte, M. R. Analysis of the interaction of Plexin-B1 and Plexin-B2 with Rnd family proteins. *PLoS One* **12**, e0185899 (2017).

411. Klein, R. M. & Higgins, P. J. A switch in RND3-RHOA signaling is critical for melanoma cell invasion following mutant-BRAF inhibition. *Mol. Cancer* **10**, 114 (2011).

412. Liu, B. *et al.* RND3 promotes Snail 1 protein degradation and inhibits glioblastoma cell migration and invasion. *Oncotarget* 7, 82411–82423 (2016).

413. Belgiovine, C. *et al.* Reduced expression of the ROCK inhibitor Rnd3 is associated with increased invasiveness and metastatic potential in mesenchymal tumor cells. *PLoS One* **5**, e14154–e14154 (2010).

414. Xia, H. *et al.* Suppression of RND3 activity by AES downregulation promotes cancer cell proliferation and invasion. *Int. J. Mol. Med.* **31**, 1081–1086 (2013).

415. McColl, B., Garg, R., Riou, P., Riento, K. & Ridley, A. J. Rnd3-induced cell rounding requires interaction with Plexin-B2. *J. Cell Sci.* **129**, 4046 LP – 4056 (2016).

416. McClatchey, A. I. Merlin and ERM proteins: Unappreciated roles in cancer development? *Nat. Rev. Cancer* **3**, 877–883 (2003).

417. De, P., Aske, J. C. & Dey, N. RAC1 Takes the Lead in Solid Tumors. *Cells* **8**, 382 (2019).

418. Zhou, K. *et al.* RAC1-GTP promotes epithelial-mesenchymal transition and invasion of colorectal cancer by activation of STAT3. *Lab. Invest.* **98**, 989–998 (2018).

419. Hein, A. L. *et al.* RAC1 GTPase promotes the survival of breast cancer cells in response to hyper-fractionated radiation treatment. *Oncogene* **35**, 6319–6329 (2016).

420. Ji, J. *et al.* Rac1 is correlated with aggressiveness and a potential therapeutic target for gastric cancer. *Int. J. Oncol.* **46**, 1343–1353 (2015).

421. Myant, K. B. *et al.* ROS Production and NF-κB Activation Triggered by RAC1 Facilitate WNT-Driven Intestinal Stem Cell Proliferation and Colorectal Cancer Initiation. *Cell Stem Cell* **12**, 761–773 (2013).

422. De, P. *et al.* RAC1 GTP-ase signals Wnt-beta-catenin pathway mediated integrin-directed metastasis-associated tumor cell phenotypes in triple negative breast cancers. *Oncotarget* **8**, 3072–3103 (2017).

423. Holkers, M. *et al.* Adenoviral vector DNA for accurate genome editing with

engineered nucleases. *Nat. Methods* **11**, 1051–1057 (2014).

424. Maggio, I. *et al.* Adenoviral vector delivery of RNA-guided CRISPR/Cas9 nuclease complexes induces targeted mutagenesis in a diverse array of human cells. *Sci. Rep.* **4**, 5105 (2014).

425. Berntsen, N. L. *et al.* Establishment of a surgical bile duct injection technique giving direct access to the bile ducts for studies of the murine biliary tree. *Am. J. Physiol. Gastrointest. Liver Physiol.* **314**, G349–G359 (2018).